

Dear David,

Back in July you generously offered a chance to comment on the JAMA paper authored by Swartout et al., the paper that is the basis for the webinar scheduled for broadcast on Tuesday. I realize that this response comes at the last moment, but it has taken this long to do a careful, methodical analysis, one that we (Jim Hopper, Allison Tracy and myself) feel confident about. Jim and Allison have struggled with significant issues in the dataset that we have access to, and those struggles have taken most of their time.

Our analysis is restricted to only one of the two datasets referred to in the JAMA paper – the “derivation” dataset that is in the public domain. We requested access to the “validation” dataset but that access has not been granted. The “derivation” dataset are the data from which the authors’ trajectory model was created. Therefore, problems with those data, and with the assumptions and trajectory analyses performed by the authors, are problems that are fundamental to their findings.

The “executive summary” written by Dr. Allison Tracy (methodologist and senior research scientist at the Wellesley Centers for Women) is just that. She will be completing a lengthy, technical analysis of the problems associated with both the “derivation” dataset itself and with the trajectory modeling performed by the authors. We will also be consulting with other methodologists to review Dr. Tracy’s analyses and, if they have the time and interest, to conduct independent analyses of their own before reviewing her analyses.

It is our intention to pursue a peer reviewed publication of our analyses, something that will require considerably more time. Our decision to share the attached analyses now was prompted by our concern that the publication of the JAMA paper has been followed by extensive dissemination of the findings, through both professional and media outlets, including the web conference taking place tomorrow. We hope that by sharing our concerns about the findings and their interpretation by the authors, an informed, open discussion can be stimulated.

The frequency analyses performed by Dr. Jim Hopper provide a very clear window onto the discontinuity between what the JAMA authors argue and what their data actually portray. It is very clear that a majority of men who commit rape and attempted rape do so repeatedly, and that the vast majority rapes and attempted rapes are committed by serial offenders.

The authors’ argument that only a small number of these perpetrators are serial offenders rests on the following assumptions that we believe are either shaky or incorrect:

1. They define serial rape in an arbitrarily restrictive way: men who report committing rapes across more than one assessment period. Thus, a man who reports that he committed multiple rapes, say 5 rapes between September and May of a given assessment period, is, according to the JAMA authors, NOT a serial offender.

2. They arbitrarily exclude attempted rapes entirely from their analysis. Attempted rape is a serious behavior, and it is clearly very related to rape. There is no justification for excluding attempted rape.

3. They assume that any subject who reports committing multiple forms of rape during a single assessment period (e.g., both rape by force and rape by intoxication) was actually referring to the same incident. They argue that the SES does not preclude this possibility. While it is true that the SES does not preclude this possibility, they neglect to acknowledge that their assumption – that all forms of rape refer to the same incident – is just as “radical” as its opposite – that all forms of rape refer to separate incidents. They have not improved the validity of the SES questions, they have merely replaced one radical assumption with its opposite. A more rational solution is to run both analyses: assume they are all one incident; assume they are all separate incidents; and let the reader judge.

4. Having made the above assumptions, the authors then ran trajectory analyses that, as Dr. Tracy identifies, appear to be rife with problems and problematic assumptions that undermine the authors’ conclusions.

5. The authors also fail to note another possible, even likely source of artifact that would dramatically affect their trajectory analyses. The SES and its derivatives function on implicit deception. We ask subjects questions about behaviors that meet the legal definition of rape, and we hope that they will not recognize what we are asking. In a longitudinal study such as this, this implicit deception is being stretched. Do some subjects – at the second, third or fourth assessment – “catch on” to the meaning of the questions and become more guarded in their responses, refusing to disclose behaviors that they have actually committed? It is a distinct possibility.

I realize that the timing of this – the eve of the webinar – is terrible. If there was any way we could have completed this work more quickly, believe me, we would have. As it is, there is still much work to be done.

Finally, we earnestly hope that an open discussion of these issues will also help to dispel what we believe to be a false conflict between the phenomenon of serial offending and the critical need for education and prevention. The phenomenon of serial offending must certainly be reckoned with, but it in no way lessens the desperate need for comprehensive and increasingly sophisticated prevention work in the arena of sexual violence.

All the best,

David Lisak & Jim Hopper

What Does “Derivation Dataset”  
Used in Swartout et al. 2015  
Tell Us About Serial Rape?  
Simple Frequency Analyses

Analyses and Presentation  
by Jim Hopper, Ph.D.

# Orientation to Information on the Following Slides - Datasets

- Swartout et al. used two different data sets:
  - “Derivation dataset” – J. White’s, 1990-1995
  - “Validation dataset” – M. Thompson’s, 2008-2011
- The “derivation dataset” is publicly available; we have been unable to obtain the “validation dataset” from M. Thompson
- Swartout and colleagues used 3 items from Koss’s Sexual Experiences Survey

# Orientation to Information on the Following Slides – Definitions 1

- Swartout et al. used 3 items from Koss's Sexual Experiences Survey to define rape as completed rape. For derivation data:
  1. **Penile-vaginal rape:** "Have you engaged in sexual intercourse with a woman when she didn't want to by threatening or using some degree of physical force (twisting her arm, holding her down, etc.)?"
  2. **Drug/alcohol-related rape:** "Have you ever deliberately given a woman alcohol or drugs and then engaged in sexual intercourse when she didn't want to?" [Note: unclear if she was incapacitated.]
  3. **Other rape:** "Have you engaged in sex acts (oral or anal intercourse or penetration by objects other than the penis) with a woman when she didn't want by threatening or using some degree of physical force (twisting her arm, holding her down, etc.)?"
- Importantly, Swartout et al. left out data from 2 items on attempted rape (penile-vaginal and drug/alcohol-related).

# Orientation to Information on the Following Slides – Definitions 2

- Swartout et al. defined “serial rape” as self-reported commission of rape at more than one assessment period.
- However, the attrition rate in the study was very high, and Swartout et al. defined men who reported committing multiple rapes before college (i.e., 2 rapes, 3-5 rapes or more than 5 rapes) and then dropped out of the study as not serial rapists.
- **Importantly, for any period assessed, a man could report committing more than 5 rapes but not be defined as a “serial” rapist, so long as he did not report committing rape at another assessment period – even if he dropped out of the study after reporting committing multiple rapes.**

# Orientation to Information on the Following Slides – Considerations 1

- Swartout et al. caution that men completing the SES could have been referring to the same rape incident with their responses to more than one item.
- Therefore, they argue, researchers should not add up the number of rape acts reported for the different SES items on completed rape to determine how many rapes a man committed during a particular period.
- Swartout's solution to this problem – to arbitrarily assume that all rape endorsements refer to the same incident – is not more or less conservative than assuming that all rape endorsements are separate incidents.

# Orientation to Information on the Following Slides – Considerations 2

The problem of how to interpret multiple rape endorsements does not apply to items covering attempted rape. By definition, acts in which the man reports attempting but failing to commit penile-vaginal rape or drug/alcohol-related rape cannot overlap with acts in which they “succeeded” in completed rapes of those types.



# Question #1

What percentage of men who reported committing rape at any particular assessment reported committing more than one rape during that assessment period?

# Percentage of Men Reporting Completed Penile-Vaginal Rape Who Were Serial Rapists

	Public Database								
	# Single Perps	# Serial Perps	# Perps Raping > 5x	Perps as Pct of Sample	Pct Serial Perps	Pct of Perps Raping >5	Pct Serial Perps Raping >5	Sample Size	Pct Original Sample
Pre-college	6	11	2	2.00	64.7	11.8	18.2	851	
Spring Freshman	3	3	1	0.93	50.0	16.7	33.3	643	75.6
Spring Sophomore	3	4	3	1.54	57.1	42.9	75.0	454	53.3
Spring Junior	1	7	4	2.67	87.5	50.0	57.1	300	35.3
Avg Pre- thru Jr					64.8	30.3	45.9		
Avg Fresh thru Jr					64.9	36.5	55.2		
Spring Senior	0	1	0		100.0	0.0	0.0	145	17.0

# Percentage of Men Reporting Completed Drug/Alcohol-Related Rape Who Were Serial Rapists

	Public Database								
	# Single Perps	# Serial Perps	# Perps Raping > 5x	Perps as Pct of Sample	Pct Serial Perps	Pct of Perps Raping >5	Pct Serial Perps Raping >5	Sample Size	Pct Original Sample
Pre-college	17	24	4	4.86	58.5	9.8	16.7	844	
Spring Freshman	9	8	2	2.63	47.1	11.8	25.0	646	75.9
Spring Sophomore	4	5	1	1.98	55.6	11.1	20.0	455	53.5
Spring Junior	2	10	1	4.01	83.3	8.3	10.0	299	35.1
Avg Pre thru Jr					61.1	10.2	17.9		
Avg Fresh thru Jr					62.0	10.4	18.3		

# Percentage of Men Reporting Completed Other Rapes Who Were Serial Rapists

Public Database									
	# Single Perps	# Serial Perps	# Perps Raping >5x	Perps as Pct of Sample	Pct Serial Perps	Pct of Perps Raping >5	Pct Serial Perps Raping >5	Sample Size	Pct Original Sample
Pre-college	8	15	5	2.70	65.2	21.7	33.3	851	
Spring Freshman	3	4	0	1.09	57.1	0.0	0.0	642	75.4
Spring Sophomore	4	3	0	1.55	42.9	0.0	0.0	453	53.2
Spring Junior	1	7	3	2.67	87.5	37.5	42.9	300	35.3
Avg Pre thru Jr					63.2	14.8	19.0		
Avg Fresh thru Jr					62.5	12.5	14.3		
Spring Senior	1	0	0		0.0	0.0		145	17.0

# Percentage of Men Reporting Attempted and/or Completed Penile-Vaginal Rape Who Were Serial Rapists

	# Single Perps	# Serial Perps	# Perps Raping >5x	Perps as Pct of Sample	Pct Serial Perps	Pct of Perps Raping >5	Pct Serial Perps Raping >5	Sample Size	Pct Original Sample
Pre-college	7	17	3	2.84	70.8	12.5	17.6	844	
Spring Freshman	4	8	2	1.86	66.7	16.7	25.0	646	75.9
Spring Sophomore	1	8	4	1.98	88.9	44.4	50.0	455	53.5
Spring Junior	0	9	4	3.01	100.0	44.4	44.4	299	35.1
Avg Pre thru Jr					81.6	29.5	34.3		
Avg Fresh thru Jr					85.2	35.2	39.8		

# Percentage of Men Reporting Attempted and/or Completed Drug/Alcohol- Related Rape Who Were Serial Rapists

	# Single Perps	# Serial Perps	# Perps Raping > 5x	Perps as Pct of Sample	Pct Serial Perps	Pct of Perps Raping >5	Pct Serial Perps Raping >5	Sample Size	Pct Original Sample
Pre-college	18	27	6	5.34	60.0	13.3	22.2	842	
Spring Freshman	7	20	2	4.19	74.1	7.4	10.0	645	75.8
Spring Sophomore	7	12	2	4.18	63.2	10.5	16.7	455	53.5
Spring Junior	2	13	3	5.02	86.7	20.0	23.1	299	35.1
Avg Pre thru Jr					71.0	12.8	18.0		
Avg Fresh thru Jr					74.6	12.6	16.6		

## Question #2

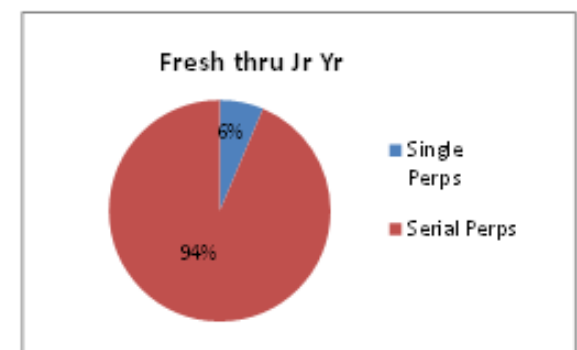
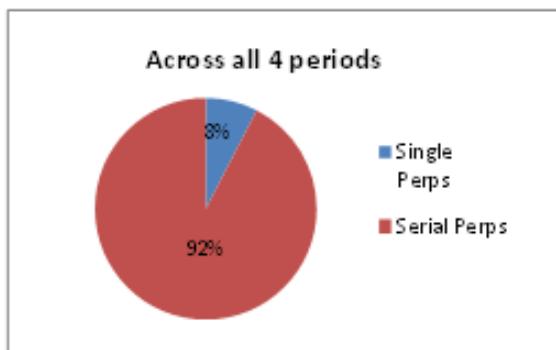
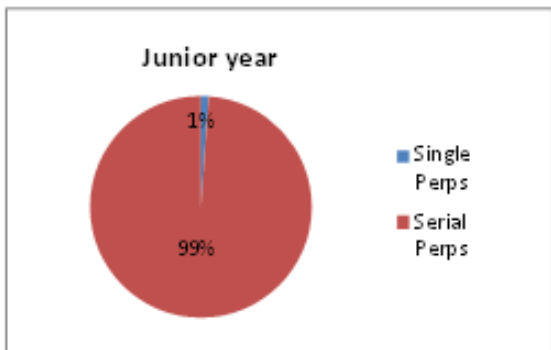
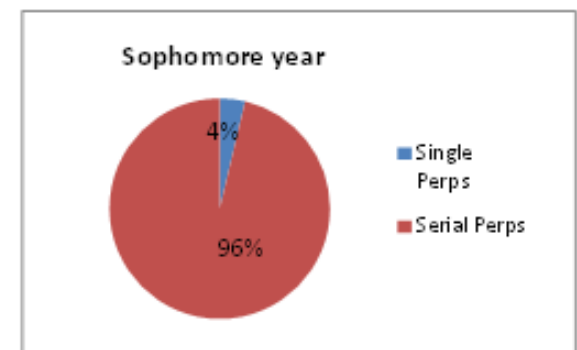
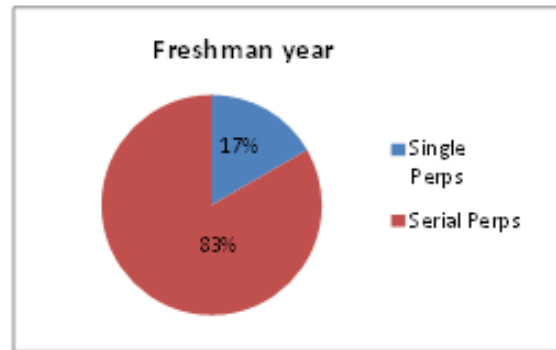
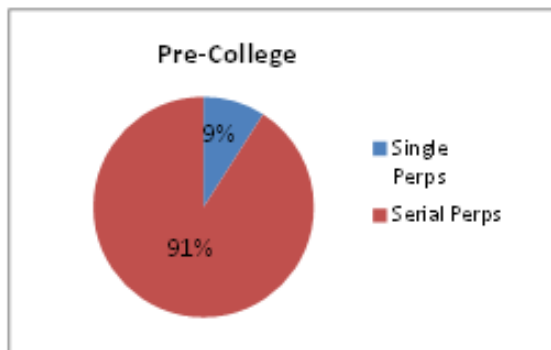
What percentage of rapes reported by all men during a particular assessment period were perpetrated by serial rapists?

# Notes on (Conservative) Calculations

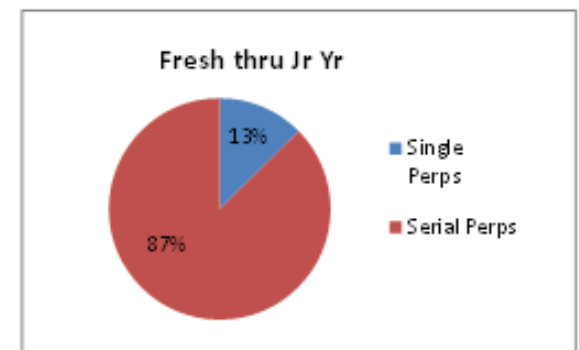
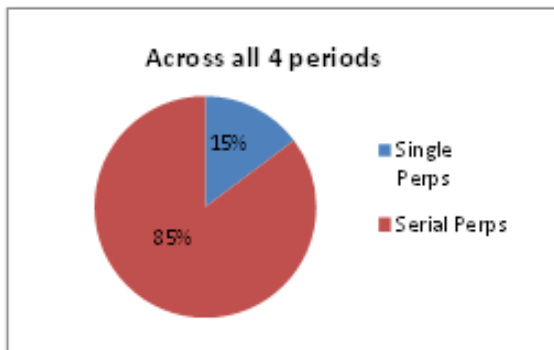
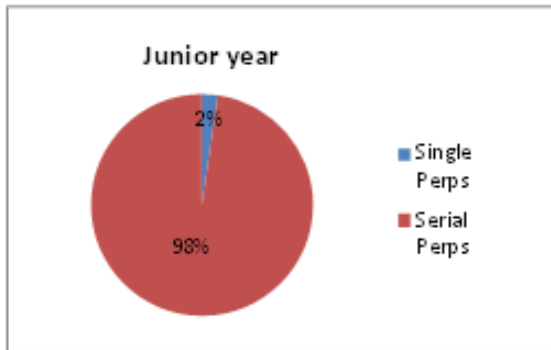
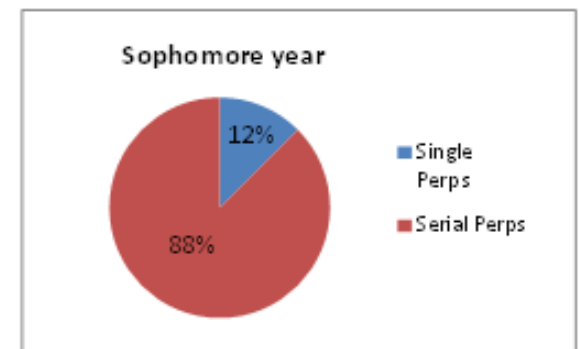
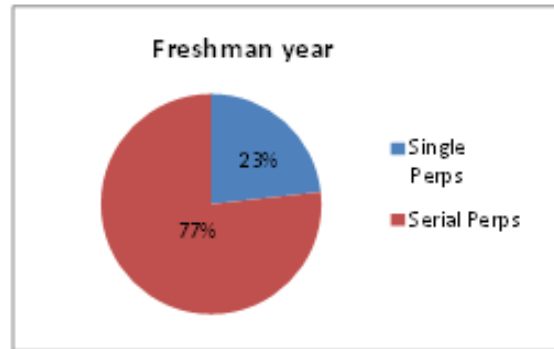
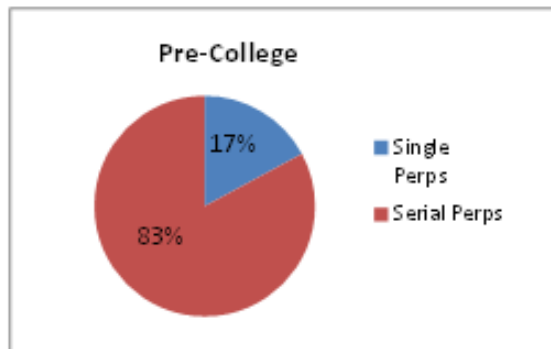
- For the following slides, all calculated total numbers of rapes (committed by serial rapists), used to generate percentages, are based on **conservatively** estimating values of 3 and 6 for, respectively, subject responses of “3 to 5” and “more than 5.”
- NO OVERLAP = All forms of rape reported by a subject in an assessment period are counted as separate incidents (e.g., 2 penile-vaginal + 2 drug/alcohol = 4; 2 penile-vaginal + 3 drug/alcohol = 5).
- 100% OVERLAP = All forms of rape reported by a subject in an assessment period are counted as a single incident (e.g., 2 penile-vaginal + 2 drug/alcohol = 2; 3 penile-vaginal + 2 drug/alcohol = 3).



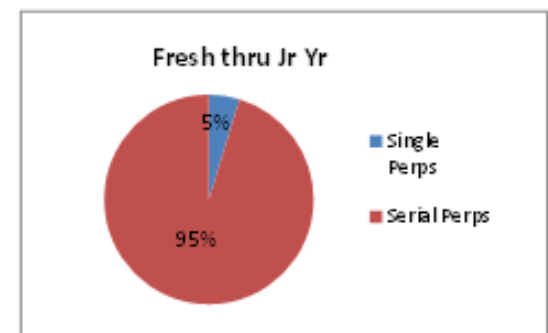
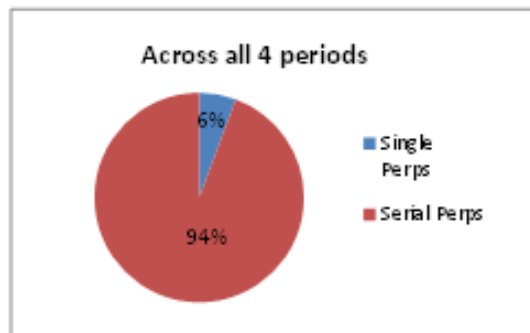
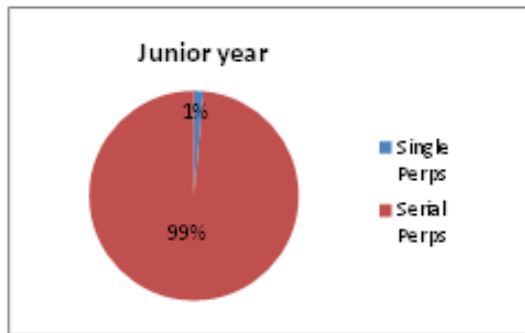
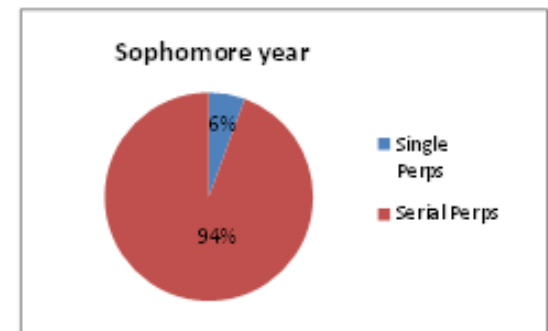
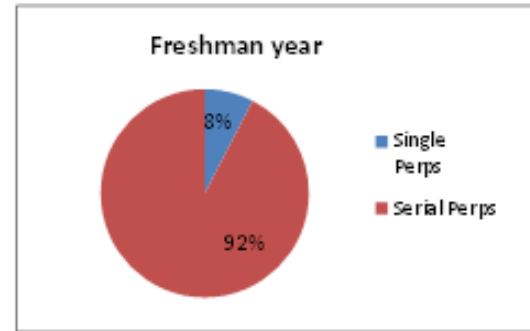
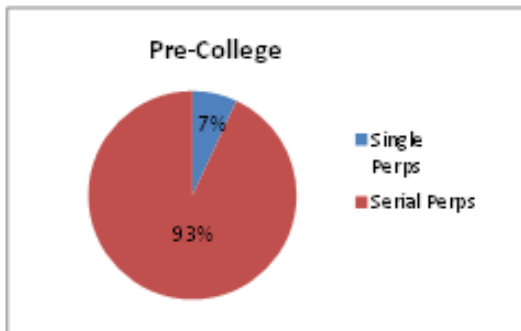
# Percentage of Completed Rapes of All Three Kinds, Assuming **No Overlap**, Committed by Serial Rapists



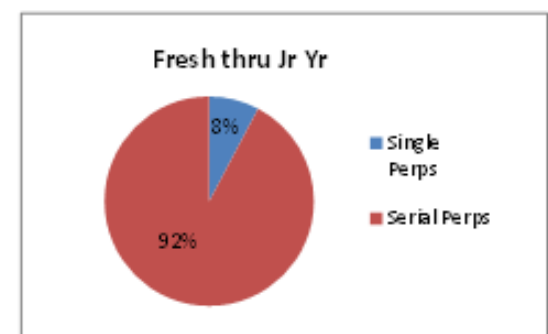
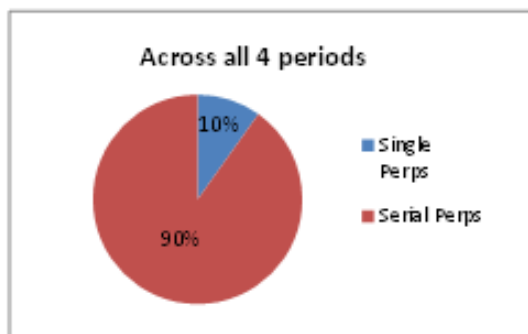
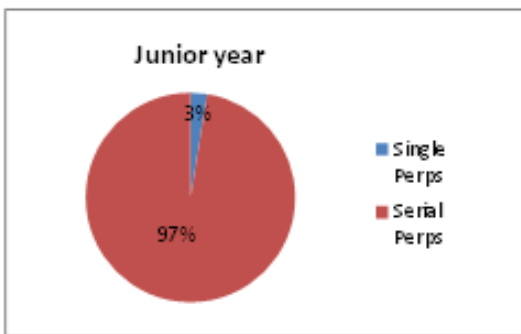
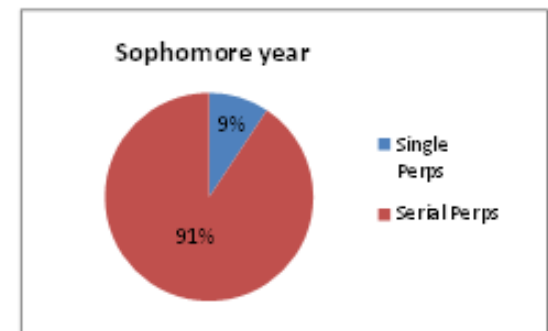
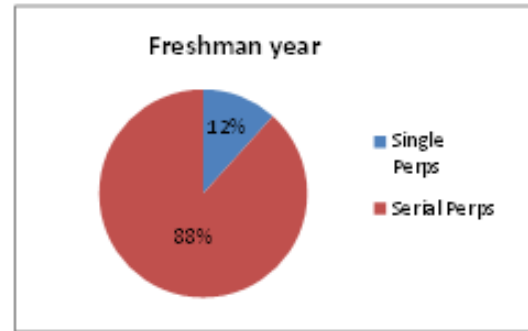
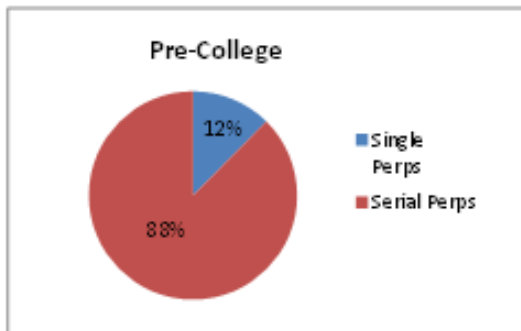
# Percentage of Completed Rapes of All Three Kinds, Assuming **100% Overlap**, Committed by Serial Rapists



# Percentage of Attempted & Completed Rapes of All Three Kinds, Assuming **No Overlap**, Committed by Serial Rapists



# Percentage of Attempted & Completed Rapes of All Three Kinds, Assuming **100% Overlap**, Committed by Serial Rapists



## Methodological Critique: Executive Summary

On the face of it, the July 13, 2015 JAMA article entitled, “Trajectory Analysis of the Campus Serial Rapist Assumption,” authored by Swartout, Koss, White, Thompson, Abbey, and Bellis appears to be scientifically sound. The paper reports a study conducted with sizeable samples from both a derivation and validation dataset to fit latent trajectories of the probability of raping across the college years, using perpetrators’ own reports. The article reads well and the naïve reader would likely be impressed and ready to accept the authors’ conclusions – that most campus rapists are *not* predatory serial rapists but rather opportunistic, time-limited rapists. However, a closer look reveals a number of problems that call the science and conclusions of the paper into question.

### *Problems with the Analyses*

With the dataset and the statistical programming code used to conduct the analyses provided by Dr. Swartout, I discovered a number of irregularities. When I ran the model for the derivation dataset, there were two warnings automatically generated by the Mplus software: (1) the model did not converge on a single best loglikelihood value and (2) the model was underidentified, so that key growth parameters were fixed rather than estimated.

Despite reporting that the fifth wave was omitted from the derivation analysis sample due to low response rates, Dr. Swartout’s programming code included the fifth wave. The consequence of using the “low response” fifth wave is that the analysis relied heavily on how the missing data were handled in the analysis. The authors claim (E3, 1<sup>st</sup> paragraph) that “...missing data were not related to reports of sexual violence across the study” citing the Pearson chi-square test for missing completely at random. A non-significant test means that the missingness did not produce biased estimates. Because there were no covariates in the analysis, this test also addressed the more central assumption of missing at random. Interestingly, when the final timepoint was *excluded* from the analysis model, the test indicated a significant departure from this central assumption (Pearson chi-square=57.319, df=39,  $p<.03$ ). In other words, the missing data handling technique used in the JAMA article may easily have resulted in biased estimates of the probability of rape.

Another thing that changed when 4 rather than 5 timepoints were used was the shape of the trajectory curves. Instead of a decreasing class, the third class was characterized by 100% probability of raping at pre-college, Year 1, and Year 3, with a 25% probability at Year 2. Patterns that yield 100% and/or 0% likelihoods often indicate that there are too few members in a given class, causing the parameter values for that class to be “overfitted” to these specific individuals. In both the 4 and 5 timepoint models, only 1% of the sample were assigned to each of the two “rapist” classes (sample sizes ranged from 6 to 12). Certainly, the statistical power for detecting such rare latent classes is very low. When sample sizes are this small, it is not surprising that the predicted probabilities are unstable, likely to change with small changes to the model. This is evident when examining the classification quality, where individuals in the two smallest classes had a significant likelihood of being misclassified (4 timepoint model: 32% & 35%; 5 timepoint model: 49% & 39%).

## *Swartout and Colleagues' Own Data Tell a Very Different Story*

It is not unusual that a very large subpopulation in a latent trajectory analysis heavily influences the solution. To focus the analysis on college-aged rapists, I omitted the very large subgroup of non-rapists (i.e., no rape reported at any timepoint) and combined the derivation and validation datasets to increase the number of rapists on which to base the trajectory analysis. Using the same syntax provided by Dr. Swartout, I arrived at a very different solution. In this model, the classes were more reasonably balanced (61%, 22%, 17%) and therefore more likely to describe prevalent subpopulations. Interestingly, only the smallest class (17%) had a pattern that described time-limited rape.

By using trajectory analysis, Dr. Swartout and his colleagues imposed a restriction on the data that the probability of rape can best be described by a smooth line over time – a severe constraint that mismodels men who rape at non-consecutive timepoints. Further, using the pre-college timepoint for which all participants have data, heavily influenced the growth trajectory – particularly the intercept term – on which the trajectory classes were based. Indeed, the “increasing” class shown graphically in the JAMA article had a 0% probability of rape at pre-college and the “decreasing” class had a 100% probability of rape at pre-college. Such heavy reliance on the pre-college timepoint, paired with the implied independence of pre-college and college rapes (E6, 1<sup>st</sup> paragraph), does not directly address rapes committed *during the college years*. When I released the constraint of a smooth trajectory, thereby allowing the model to freely estimate the probability of rape at each timepoint, I obtained a 5-class model in which 65% of the rapist sample had at least 40% probability of rape at two or more of the four timepoints.

All the above models showed the same estimation problems as did the JAMA models: the models did not converge and were underidentified. There is an alternative model, logistic autoregression, which addresses the research question in a way that avoids the computationally intensive and heavily assumption-dependent latent trajectory modeling. This approach uses rape at one timepoint to predict rape at the next timepoint. The model converged on a solution without difficulty, had ample statistical power since all participants' data were used (rapist and non-rapists), and minimized the impact of missing data since estimates depended only on consecutive timepoints rather than the full study period. The fit of this model was excellent (chi-square=1.203, df=3,  $p=.75$ , RMSEA=0.00, CFI=1.00) and each regression parameter was significant to the  $p<.001$  level. Based on this model, the predicted probability of raping during Freshman year was more than 5 times higher (OR=5.64) if a man had raped pre-college; the probability of raping during Sophomore year was almost twice as high (OR=1.95) if a man had raped during Freshman year; and the probability of raping during Junior year was more than 2 times higher (OR=2.29) if a man had raped during Sophomore year.

## *Returning to the Primary Dataset: Problems with Data Integrity and Validity*

In an effort to better understand the data and describe the men identified as rapists, I went to the original public-use dataset from which Dr. Swartout and his colleagues created their derivation dataset, the *Longitudinal Study of Violence Against Women: Victimization and Perpetration Among College Students in a State-Supported University in the United States, 1990-1995* (ICPSR 3212). Unfortunately, the original case ID numbers were absent in the Mplus analysis dataset used for the article and, despite multiple requests, no file was provided to match the JAMA dataset to the public-use

dataset. Descriptively, the raw frequencies for each of three sexual assault items used by the JAMA study to operationalize rape clearly showed that the majority of college rapists rape multiple times within timepoint as well as across timepoints.

Using the data management programming code exactly as it was provided by Dr. Swartout, I reconstructed the dichotomous rape variables (i.e., variables indicating yes vs. no for rape at particular timepoints) and reran the JAMA analysis. The reconstructed analysis did not replicate the results presented in the article, and the frequencies of the rape variable did not match the frequencies of the variables in the JAMA analysis dataset. In an attempt to understand this disconnect, I examined the variables on which the JAMA rape dichotomies were based. Oddly, Swartout's code did not draw directly from the raw frequency data that the respondents provided but instead drew from recoded dichotomies of each of three sexual assault items operationalizing rape. Importantly, a cross tabulation of the dichotomized sexual assault indicators with the raw data showed that in some cases, missing data were assigned as "no rape." This has the effect of underestimating the rapes that men in the sample committed.

There was a great deal of missing sexual assault data in the original dataset and the proportion of missing data in the original dataset did not match the missing data in the JAMA analysis dataset (JAMA : 25%, 47%, & 65%; Original: 25%, 60%, & 73% at Times 2, 3, & 4, respectively). In addition to study attrition, there was a substantial amount of unexplained missingness where men participating in the study (completing at least some of the survey) did not provide data for any of the sexual assault indicators (within missing data: 18%, 13%, & 29% at Times 2, 3, & 4, respectively). The extent of missing data, the unexplained missing data patterns, and the missing data miscoded as no rape are serious issues that undermine the validity of the analysis and the conclusions reported in the JAMA article.

### *Choices that Underestimate Serial Rape and Serial Rapists*

In addition to the miscoded missing values, several other decisions and actions by the authors resulted in underestimation of the numbers of serial rapes and serial rapists. For example, by creating a single dichotomous indicator of rape for each timepoint, the authors ignored multiple rapes within timepoint, thereby underestimating and "defining away" serial rapists (e.g., men who reported committing "more than 5" rapes freshman year, but then dropped out of the study, were not defined as serial rapists). By assigning a single rape indicator to multiple reports of rape across sexual assault items, the authors assumed that all responses to rape items refer to the same rape incident, again underestimating serial rapists. Finally, the authors underestimated serial rapists by excluding attempted rapes.

### *Conclusion*

The scientific integrity of the study described in the JAMA article is highly suspect at best. As it stands, the article contains erroneously coded data, misalignment with the raw data, analyses that are based on untenable model assumptions and fail to meet criteria of validity, and underestimate the prevalence of serial rapists in their sample and the percentage of rapes those serial rapists report committing. No reasonable and scientifically grounded debate over the "serial (campus) rapist hypothesis" can depend on this study.