Senior Seminar

Criminology in Crisis? Replication and Reproducibility in Criminology and the Social Sciences

CRM 495 – 002

T/TH 9:30 – 10:45am in Bear Hall, Rm. 281

Spring, 2022

Professor: Jake Day, PhD
Email: dayj@uncw.edu
Office Hours: Thursday 1-3pm

Office: Bear Hall, Rm. 224
Office Phone: 910-962-7024

Course Introduction

What do we "know" in Criminology? Ask a professional criminologist or criminal justice practitioner and you'll likely get varied answers, ranging from a specific study's research findings to general statements about research evidence, usually prefaced with some version of "research says…". Indeed, at this point in your undergraduate studies, your prior coursework has likely exposed you to the criminological "canon," or those phenomena and ways of studying those phenomena that working criminologists take for granted as "established methods and facts."

In this course, we will question that "canon" with the specific aim of critically assessing "what we know" in Criminology. In doing so, we will discuss challenges encountered when attempting to produce trustworthy and reliable empirical findings and identify some strategies and build some practical skills for overcoming such challenges.

These issues have garnered increasing attention over the last decade alongside the emergence of the well-documented "replication crisis" in (social) science. There seems to be a growing recognition (among some) that many of the issues identified as causes of the "replication crisis" in other disciplines also are present – and perhaps even highly prevalent – in the day-to-day practice of scientific criminology and criminal justice. Yet, as disciplines, criminology and criminal justice have been relatively slow in adopting reforms meant to increase veracity, reliability, and trust in our empirical findings.

I jointly designed this course in collaboration with Dr. Jonathan R. Brauer (Indiana University), a former graduate student colleague, current collaborator, and long-time friend. Together, we initially designed the course as a first-year methods course for graduate students. Here, I have redesigned the course specifically for senior-level undergraduates in Criminology and Criminal Justice.

I expect that knowledge of and the technical skills necessary to practice open and reproducible science will become increasingly valued in professional criminology and criminal justice settings. I would rather you start your professional careers with some of these knowledge and skills as opposed to retooling later like I am doing. I also do not want you to graduate with an overly naïve or overly cynical perspective regarding "what we know" in Criminology and Criminal Justice. Rather, I think you are better off with a realistic and critical view of the methodological challenges that are perhaps endemic to social scientific fields. You cannot start to imagine changing typical practices and/or addressing these challenges until you have a clear picture of what you are facing.

At minimum, my hope is that through this course you develop an appreciation for the scientific enterprise, the importance of research design, and the current debates and reforms related to practicing open and reproducible social science and its role in criminological research and practice. I hope you will also come to see that learning how to collect, analyze, and interpret data is just that – an ongoing process of *learning*. Likewise, as a professional criminologist or criminal justice practitioner, I hope you leave this course prepared to continue the difficult but rewarding investment in this learning process throughout your formal education the rest of your professional career.

Course goals and objectives

My primary goal in this course is to guide and evaluate your pursuit of the following objectives:

- *Understand meaning, causes, and possible solutions to the so-called "replication crisis."*
- *Relatedly, identify various challenges encountered when attempting to produce trustworthy and reliable empirical findings in social science research, as well as identify potential solutions to these challenges.*
- *Develop a capacity to review, evaluate, and apply social science research methods and to critically assess "what we know."*
- *Develop an appreciation for, and practice, some of the technical skills necessary to conduct one's own open, reproducible, and trustworthy social science research.*

In order to accomplish these primary goals, you will be expected to do the following:

- *Read assigned materials, then summarize, react, and ask questions about the readings.*
- *Demonstrate engagement with and comprehension of course materials by being prepared for and participating in class discussions in an intelligent and informed way.*
- *Complete assignments intended to provide a basic introduction in how to conduct reproducible social science research using R and Rstudio.*
- *Reproduce or conceptually replicate an existing research study following reproducibility principles using R and conduct a peer review of a classmate's reproducibility project.*

My Assumptions about Students' Knowledge and Skills

If you are in this class, I assume you have completed the core classes in the Criminology major sequence, particularly SOC/CRM 255: Criminology, SOC/CRM 300: Social Research Methods, and SOC/CRM 301: Data Analysis. This means I expect you to have some basic knowledge of the commonly held "causes and correlates" of crime, the fundamentals of social research design (e.g., sampling and measurement) and data analysis (e.g., characteristics of a quantitative data set and descriptive measures of central tendency and dispersion like the mean and standard deviation).

I ***do not*** assume you have any background in advanced statistical computing, including wrangling and analyzing data by writing code in a statistical programming environment like R. Of course, I recognize that some of you may have this experience and it should serve you well throughout the course. Assignments in this course are designed to provide you with a basic introduction to conducting reproducible data science using R and Rstudio.

For those of you who feel like you need a refresher on any of these topics or want to dive into more details about open-source statistical computing, supplementary resources will be provided on Canvas as well as in the instructions for specific assignments.

Required Text

Ritchie, Stuart. 2020. *Science Fictions*. New York: Metropolitan Books.

*Additional readings will be made available on the course's Canvas site

Course Requirements

Although obtaining a good grade in the class should ***not*** be your main goal (your goal should be to prepare yourself for professional work, which really has little connection with your grades in college), your performance on the following activities and assignments will contribute to my overall evaluation of your performance in the following proportions:

1) **Course Participation & Contribution (5%)**

   Each student will be expected to read and study carefully all the assigned material prior to the class meeting in which it is to be discussed, and to participate effectively in the class discussion. It is acceptable to be wrong, to misinterpret, to challenge/disagree, or simply to misunderstand. It is not acceptable to be unprepared.

2) **Reading Assignments (20%)**

   For each class in which you have assigned readings (marked with asterisk next to the date on the schedule), you are required to submit a "Reading Assignments" document with the following sections (details on Canvas):

   - Reading Summary – Summarize the readings in a few sentences.
   - Reading Reaction – Describe reactions you had to readings in a few sentences.
   - Discussion Questions – Write a discussion question based on the readings.

   Reading Assignments should be submitted via Canvas by 8pm on the Monday night before the reading is to be discussed in class on Tuesday. In addition to assessing whether Reading Assignments were turned in on time, I will evaluate these summaries, reactions, and questions for diligence, thoughtfulness, pertinence to the weekly topic of discussion, and efforts at comprehension displayed. There are 14 reading assignments throughout the semester. For attendance flexibility purposes, I will only count 12 reading assignments toward your final grade.

   Note: Any additional assignments not listed on the syllabus will count toward this portion of your grade.

3) **R Assignments (25%)**

   Regular assignments will provide you with a basic introduction to conducting reproducible data science using R and RStudio (see Canvas for details). After completing these assignments, you should have a firm foundation for reproducing results from an existing study (see "Replication & Reproducibility Project" below).

   - R Assignment 1: Introduction to R & RStudio
   - R Assignment 2: Introduction to RMarkdown
   - R Assignment 3: Data Visualization
   - R Assignment 4: Reproducible File Structure
   - R Assignment 5: Downloading & Describing Data
   - R Assignment 6: Data Wrangling
   - R Assignment 7: Building Publishable Tables

   To receive full credit for the assignment, you must make an earnest attempt and submit your assignment by 9:30am on the Thursday it is due. After we discuss and workshop issues with the assignment in class, you will have until 9:30am on Friday to make any revisions necessary and submit an updated file to Canvas and to your OneDrive folder.

4) Replication & Reproducibility Project (50%)

There are seven phases to the *Replication & Reproducibility Project*, listed as a "RR Project Phase (1–7)" on the course schedule below. To complete this project in its entirety, you will likely need to combine what you learned in the "R Assignments" (above) with additional self-directed learning via provided materials on the course Canvas page as well as online searches for additional R vignettes; this is a normal part of data science – in fact, it something that I do on an almost daily basis.

- Phase 1: Familiarize yourself with the NYS Data and identify a potential research topic/question.
- Phase 2: Identify published research on potential research topic/question and find an existing article using NYS Data.
- Phase 3: Describe & Justify Reproduction or Conceptual Replication.
- Phase 4: Identify and describe key variables
- Phase 5: Submit First Draft of Replication and Reproducibility Project
- Phase 6: Peer Review
- Phase 7: Final Replication and Reproducibility Project

Grade Scale

A : 93-100%	B+ : 87-89%	C+ : 77-79%	D+ : 67-69%	F : < 60%
A- : 90-92%	B : 83-86%	C : 73-76%	D : 63-66%	
	B- : 80-82%	C- : 70-72%	D- : 60-62%	

Classroom Environment & Format

This course is meant to be a seminar. I would prefer to moderate an active discussion regarding key issues surrounding replication and reproducibility issues within criminology and criminal justice rather than lecture about these topics. This style of learning can be both intellectually challenging and rewarding. However, you should recognize that in a seminar, everyone is both a student and a teacher and we all share the responsibility for creating a successful learning environment. This includes being on time, prepared for class and participating in classroom discussion and activities.

I expect that discussions will be professional and polite but engaged. Do not shy away from making points, asking pointed questions, or pushing arguments about the topic(s) of the day. I will try to push you on your arguments and ideas, and I trust that you will do the same for me and each other.

Technical Issues

Technical and logistical problems, such as being unable to access a computer, computer failure, problems with internet connections (such as speed or quality of the connection) or browser, failure to check that your assignments have properly uploaded, etc., will not automatically result in remedies favorable to students. Even if the technical or logistical problem is not your fault, you are not guaranteed a retake or "do-over" for the assignment. Any such issues are dealt with on a case-by-case basis. Further, to avoid last minute problems, it is highly recommended that you complete readings, videos, and assignments as early as possible during each module.

Students with Disabilities
UNCW is open and accessible to students with disabilities and is committed to providing assistance to enable qualified students to accomplish their educational goals, as well as assuring equal opportunity to derive all of the benefits of campus life. If you are a person with a disability and anticipate needing accommodations of any type in order to participate in this class, you must notify the Disability Resource Center (#1033 DePaolo Hall, 910-962–7555), provide necessary documentation of the disability and arrange for the appropriate authorized accommodations. Once these accommodations are approved, please identify yourself to me so we can implement these accommodations.

Harassment
The role of all employees and students is to create and maintain a supportive and harassment--free working and learning environment for all members of the campus community. All faculty, staff and students are responsible for understanding and complying with harassment policies. These policies can be viewed at: http://uncw.edu/policies/documents/02.200_Unlawful_Harassment.050605.pdf.

> ***Title IX Mandatory Reporting Policy***
> UNCW takes all forms of sexual harassment and sexual misconduct very seriously. When students disclose, first or third-hand, to faculty or staff about sexual harassment or misconduct, this information must be reported to the administration for follow-up. The purpose of this is to insure that students' rights are insured, appropriate resources are offered, and further investigation is explored. Three offices are confidential, and thus do not need to make this report: UNCW CARE, the Student Health Center, and the Counseling Center. If you want to speak to someone in confidence, these resources are available, including CARE's 24-hour crisis line (910-512-4821). For more information, please visit www.uncw.edu/sexualmisconduct or www.uncw.edu/care.

Academic Integrity
Academic integrity is an important issue in higher education and will be treated as such in this class. Graduate students are expected to be familiar with the basic standards of citing others' work, giving credit for others' ideas, and avoiding plagiarism. As an enrolled student, you are agreeing to abide by UNCW's Student Academic Honor Code (http://www.uncw.edu/odos/honorcode/).

Campus Student Learning Resources

The University Learning Center (www.uncw.edu/ulc), DePaolo Hall 1056 & 1003
The University Learning Center's (ULC) mission is to help students become successful, independent learners. Tutoring at the ULC is NOT remediation: the ULC offers a different type of learning opportunity for those students who want to increase the quality of their education. ULC services are free to all UNCW students and include the following:

- Learning Services (University Studies) http://www.uncw.edu/ulc/learning/index.html
- Math Services http://www.uncw.edu/ulc/math/index.html
- Study Sessions http://www.uncw.edu/ulc/includes/StudySessions.html
- Supplemental Instruction http://www.uncw.edu/ulc/si/index.html
- Writing Services http://www.uncw.edu/ulc/writing/index.html

Randall Library (http://library.uncw.edu)
The Randall Library has a host of online and in-person tools and services to assist students throughout their time at UNCW. This includes research services that are relevant to your work in CRM/SOC 255 (http://library.uncw.edu/get_started).

Tentative Course Schedule

** indicates reading assignment is due by 8pm the night before (Monday).*

	1: Introduction	
Week 1		
1/13 (TH)	Get acquainted; Explain course logic, schedule, & assignments	
	Optional readings:	
	Pacheco-Vega:	Preparing for reading-intensive seminar & AIC method
	2: Science: A Candle in the Dark or a House of Cards	
Week 2		
1/18 (T)*	Ritchie (p.1-24)	*Preface & Chapter 1: How Science Works*
1/20 (TH)	**R Assignment 1**	***Introduction to R & RStudio***
	Optional Videos:	
	Navarro (2021)	Running R on your own computer
Week 3		
1/25 (T)*	Ritchie (p.25-43):	*Chapter 2: The Replication Crisis*
	Optional readings:	
	Bem (2011)	*Feeling the Future*
	Open Science Collab. (2015)	*Estimating the Reproducibility of Psychological Science*
1/27 (TH)	**R Assignment 2**	***Introduction to R Markdown***
	Optional videos:	
	Navarro (2021)	*Starting R Markdown*
	3: Social Norms & Deviance in Science *	
Week 4		
2/1 (T)*	Ritchie (p.48-80):	*Chapter 3: Fraud*
	Optional readings:	
	Janz & Freese 2021	*Replication Golden Rule*
2/3 (TH)	**RR Project – Phase 1**	***Review NYS Data & Develop Research Topics/Questions***
	Optional readings:	
	Elliott (1985)	*Description of NYS Sampling and Delinquency Measures*
Week 5		
2/8 (T)*	Bartlett (2019)	*Stewart Retractions*
	Pickett (2020)	*Stewart Retractions*
	Optional readings:	
	Neyfakh (2015)	*Alice Goffman & Ethnics of Ethnography*
2/10 (TH)	**R Assignment 3**	***Data Visualization***
	Optional videos:	
	Navarro (2021)	*Data visualization with ggplot*

	4: Is Scientific "Knowledge" Full of Bias?	
Week 6		
2/15 (T)*	Ritchie (p.81-122):	*Chapter 4: Bias*
2/17 (TH)	**RR Project – Phase 2**	***Identify Published Research on Topic/Question***
Week 7		
2/22 (T)*	Roscigno & Preito-Hodge (2021)	*Racist Cops?*
	Peyton (2021)	*Racist Cops?*
	Optional Readings:	
	Hu (2021)	*Race, Policing, & Limits of Social Science*
2/24 (TH)	**R Assignment 4**	***Reproducible File Structure***
	Optional videos:	
	Navarro (2021)	*Project structure*
	Long (2020)	*A Computational Workflow for Reproducible Results*
	5: Negligence in Social Science	
Week 8		
3/1 (T)*	Ritchie (p.123-144):	*Chapter 5: Negligence*
3/3 (TH)	**RR Project – Phase 3**	***Describe & Justify Reproduction or Replication***
Week 9	**Spring Break! (March 5th – 13th)**	
Week 10		
3/15 (T)*	Knox & Mummolo (2020)	*Race & Officer-Involved Shootings*
	Johnson et al. (2020)	*PNAS Article Retraction*
	Massey & Waters (2020)	*PNAS Article Retraction*
3/17 (TH)	**R Assignment 5**	***Downloading and Describing Data***
	Optional videos:	
	Navarro (2021)	*Say hello to your data*
	Optional Readings:	
	Kastellec & Leoni (2007)	*Graphs Not Tables*
	6: Don't Believe the Hype!	
Week 11		
3/22 (T)*	Ritchie (p. 145 – 172)	*Chapter 6: Hype*
	Optional Readings:	
	Singal (2020)	*Quick Fix Psychology*
3/24 (TH)	**RR Project – Phase 4**	***Identify & Describe Key Variables***
	Optional videos:	
	Navarro (2021)	*Say hello to your data*

Week 12		
3/29 (T)*	Miller et al. (2020)	*Police & Bias Training*
	Kaste (2020)	*Police & Bias Training*
3/31 (TH)	**R Assignment 6**	***Data Wrangling***
	Optional videos:	
	Navarro (2021)	*dplyr, or a dance with data*
	7: Perverse Incentives Spoil Science	
Week 13		
4/5 (T)*	Ritchie (p. 175 – 198):	*Chapter 7: Perverse Incentives*
4/7 (TH)	**R Assignment 7**	***Building Publishable Tables***
Week 14		
4/12 (T)*	Gartner et al. (2012)	*Salami Slicing*
	Cullen et al. (2013)	*Salami Slicing*
	Optional Readings:	
	Cohen (2012)	*Overly Similar Publications*
4/14 (TH)	NO CLASS	
	8: Questionable Research Practices (QRPs) & Open Science Practices (OSPs)	
Week 15		
4/19 (T)*	Ritchie (p.199-237):	*Chapter 8: Fixing Science*
	Optional Readings:	
	Ritchie (p.239-254):	*Epilogue & Reading a Scientific Paper*
4/21 (TH)	**RR Project – Phase 5**	***First Draft of RR Project***
Week 16		
4/26 (T)*	Chin et al. (2021)	*QRPs & OSPs in Criminology*
	Optional Readings:	
	Navarro (2020)	*Paths in Strange Spaces*
4/28 (TH)	**RR Project – Phase 6**	***Peer Review***
Week 17		
5/5 (TH)	**RR Project – Phase 7**	***Final Replication and Reproducibility Project***

Replication & Reproducibility Project
(50% of your course grade)

Project Overview:

Over the course of the semester, you will be responsible for developing a research project that directly reproduces or conceptually replicates the results of a published empirical study using the National Youth Survey (NYS) Data.[1] The primary purpose of the project is to give you the opportunity to apply many of the analytical skills you have learned throughout your time at UNCW as well those you will develop throughout this course.

This project is designed to cement your appreciation for and guide you through the process of producing reproducible research. Throughout the course, you will read about extent issues and problems in the replicability and reproducibility of social science research, including in Criminology and Criminal Justice research. But to truly develop an appreciation for the challenges researchers face in attempting to develop reproducible and replicable research, I think you need to get your hands dirty with real-world data. You will have the opportunity to do just that with this project.

Specifically, your replication and reproducibility project will be a version of one of the following options:

1) ***Direct Reproduction*** – reproduce or verify an original study's findings using the same data and methods (e.g., coding decisions).
2) ***Conceptual Replication –*** test the repeatability, robustness, or generalizability of a theoretical or observational claim from a previous study using new data and/or new measurement procedures that are conceptually similar but not identical to those used in the previous study.

Note that the difficulty of this project varies both between and within these two options. Conducting a conceptual replication will generally be more difficult than a direct reproduction. Also, while you will only be required to reproduce or replicate and (re)interpret the ***descriptive findings*** of an existing study, reproducing or replicating more results, including the ***multivariate findings*** (e.g., bivariate correlations, multiple regression, etc.) is considered more difficult. Also, characteristics of the study you choose to reproduce or replicate may also lead to more difficulty (e.g., a study with more variables, using multiple waves of data, and where more data wrangling is required). Credit (your grade) will recognize the difficulty of the task. However, it is better to do a less difficult task well than to do a more difficult task poorly. Choose wisely what you will attempt.

Although the final replication and reproducibility project will be due at the end of the semester, I will require you to complete preliminary projects and drafts for developing the project throughout the semester. This will ensure that you spend adequate time on each phase of the research and writing process involved in this project. It will also give you a chance to get feedback from me and your fellow students before the final project is due. Below is a summary of each phase of the project, but more detailed instructions will be provided for each phase on the course Canvas page.

[1] In special circumstances, you may use a different data set than the NYS. However, you should be aware that using a different data set will require you to become an expert on that data set with minimal support from the instructor and your classmates. One of the reasons for limiting the project to the NYS data is because it will allow us the opportunity to learn from each other (e.g., one student's issue with the data will be more likely to generalize to other students' projects). You must have a convincing argument and get approval form the instructor to use data other than the NYS.

Project Phases:

Phase 1: Review NYS Data and Develop Research Topics/Questions (25 points)

- In this phase, you will familiarize yourself with the NYS Data that we will be using throughout the course and identify potential research topics and/or questions that you could potentially examine with these data.

Phase 2: Identify Published Research on Topic/Question (50 points)

- In this phase you will:
    1. Find five (5) empirical research articles related to the topic or question you identified in Phase 1. You will evaluate these five articles in terms of their adherence to basic open science practices (e.g., data availability, replication code availability, etc.).
    2. Second, you will identify an empirical research article related to your topic or question that analyzes data from the NYS and review its characteristics.
    3. Third, you will tentatively commit to performing a conceptual replication of an article that does not use the NYS or a direct reproduction of one that does and identify what aspects of the research article you plan to conceptually replicate or reproduce.

Phase 3: Describe and Justify Reproduction or Conceptual Replication (50 points)

- Based on feedback from the instructor on Phase 2, here you will fully commit to a direct reproduction or conceptual replication using the NYS data. Specifically, you will provide:
    a. A description of the article, data source, and specific findings that will be replicated or reproduced, along with a justification for the reproduction. Think of it like the introductory section of a published replication/reproduction article that must describe the original research and justify the replication/reproduction research.
    b. An image of the table(s) and/or figure(s) found in the original study that you plan to reproduce.

Phase 4: Identify and Describe key Variables (50 points)

- Here you will begin working on the reproduction or conceptual replication itself. This will involve:
    1. Downloading and reading the data into R
    2. Identifying and describing all the variables (i.e., "key variables") needed to reproduce or replicate the original article's table(s) or figure(s)
    3. Summarizing the raw versions of the key variables or items.

Phase 5: Submit First Draft of Replication and Reproducibility Project (100 points)

- In this phase, you will complete a rough draft of your replication and reproducibility project. At minimum, this will require that you:
    1. Wrangle and recode all of the variables needed for reproduction.
    2. Create polished and publication-ready versions of descriptive table(s) and/or figure(s) you specified in earlier phases of the project.
    3. Interpret or reinterpret the results of your reproduction or conceptual replication independently and in relation to the original article.
    4. Describe any issues you encountered with reproducing or replicating the results of the published article.

Phase 6: Peer Review (100 points)

- Immediately following submission of the first draft of your project, I will send the link your submitted project's shared OneDrive folder to ONE of your classmates to review. I will loosely follow a "single-blind" review format. That is, the reviewer will know the author, but I will not formally identify the reviewer's name to the author. Of course, in small classes, reviewer anonymity cannot be guaranteed.
- ***Note:*** You must submit the first draft of your project on time (i.e., by **5pm on Thursday, 4/21**) to participate and earn points in this required peer review. Failure to submit the project on time may result in a "0" on this assignment.
- Peer reviewers are responsible for providing detailed and constructive feedback (i.e., not just "good job!" – there are always ways we can improve our work) using a helpful and professional tone. More detailed guidance for what to look for in your peer review will be provided on the course's Canvas page.

Phase 7: Final Replication and Reproducibility Project (125 points)
Based on feedback from your peer reviewer and instructor, submit a revised version of your replication and reproducibility project. This final version should be as polished as possible (e.g., free of grammatical, spelling, and coding errors). The computational goal is for your work to be completely reproducible (e.g., "one-click reproducibility") while the substantive goal should be to write a well-organized, thorough, and thoughtful "article" replicating or reproducing an existing study