2017

Grave Crimes and Weak Evidence: Fact-Finding Evolution in International Criminal Law

Nancy Amoury Combs

*William & Mary Law School*, ncombs@wm.edu

---

Repository Citation


http://scholarship.law.wm.edu/facpubs/1867

Copyright c 2017 by the authors. This article is brought to you by the William & Mary Law School Scholarship Repository.

http://scholarship.law.wm.edu/facpubs
Grave Crimes and Weak Evidence: A Fact-finding Evolution in International Criminal Law

Nancy Amoury Combs*

International criminal courts carry out some of the most important work that a legal system can conduct: prosecuting those who have visited death and destruction on millions. Despite the significance of their work—or perhaps because of it—international courts face tremendous challenges. Chief among them is accurate fact-finding. With alarming regularity, international criminal trials feature inconsistent, vague, and sometimes false testimony that renders judges unable to assess with any measure of certainty who did what to whom in the context of a mass atrocity. This Article provides the first-ever empirical study quantifying fact-finding in an international criminal court. The study shines a spotlight both on the testimonial deficiencies that impede accurate fact-finding and on the judges' assessments of deficient witness testimony. Although my previous work on fact-finding has been generally critical of international criminal courts, this large-scale empirical study provides far more reason for optimism. This study reveals a host of interesting and sometimes unexpected findings. Taken as a whole, however, it depicts a criminal justice system that labors in the face of severe fact-finding challenges but that has, over the years, appropriately altered its fact-finding practices to respond to those challenges.

I. INTRODUCTION

International courts that prosecute crimes such as genocide, war crimes, and crimes against humanity have provided human rights advocates with a novel and potentially powerful enforcement tool. These courts have achieved many notable successes in the twenty years since the first modern tribunal was created. The International Criminal Tribunal for Rwanda (ICTR), for instance, has played a vital role in developing the law of genocide and the prohibition of sexual violence in international criminal law. The International Criminal Tribunal for the former Yugoslavia (ICTY) is considered the first international criminal tribunal of the modern era.

* Ernest W. Goodrich Professor of Law, William & Mary Law School. I owe tremendous thanks to Eric Kades, who performed the statistical analyses described in this paper and also provided extraordinarily helpful editorial comments. For early feedback, I am grateful to Kristen Boon, Evan Criddle, Shahram Dana, Ashley Deeks, Meg deGuzman, Adil Haque, Saira Mohamed, Yvonne McDermott Rees, Daryl Robinson, James Stewart, Jenia Turner, and Alex Whiting. Finally, this Article benefited from excellent research assistance provided by Mary Antley, Daniel Biegler, James Damon, Sarah Deutch, and Caroline Lochabay. Any errors are my own.

1. The International Criminal Tribunal for the former Yugoslavia (ICTY) is considered the first international criminal tribunal of the modern era.


tional Criminal Tribunal for the former Yugoslavia (ICTY) and the Special Court for Sierra Leone (SCSL), for their parts, made history by indicting sitting heads of state,4 whereas the SCSL and the International Criminal Court (ICC) have issued ground-breaking decisions involving the enlistment and conscription of child soldiers.5 In addition to these successes, however, international criminal courts and tribunals have also faced a multitude of challenges. International criminal tribunals must contend with obstructionist defendants who seek to delay and distort proceedings,6 insufficient budgets that require them to pit fairness against efficiency,7 and recalcitrant governments that refuse to surrender indictees,8 evidence,9 and witnesses.10

Additionally, and equally pressingly, international criminal tribunals must contend with a host of factors that impede their ability to find accurate facts. Many of these factors have been on embarrassing display lately at the

5. See, e.g., Prosecutor v. Norman, Case No. SCSL-04-14-AR72(E), Decision on Preliminary Motion Based on Lack of Jurisdiction (May 31, 2004); Prosecutor v. Lubanga, ICC-01/04/01/06-2842, Judgment pursuant to Art. 74 of the Statute, paras. 478–84 (Mar. 14, 2012), https://www.icc-cpi.int/Cour
ICCs. There, judges have already refused to confirm charges against several suspects,\textsuperscript{11} have acquitted one of the court’s first four defendants,\textsuperscript{12} and have sharply scolded prosecutors for the evidentiary deficiencies in their cases.\textsuperscript{13} Fact-finding impediments have never been a more serious problem in international criminal law.

In previous works, I brought to light a wide range of deficiencies common to eyewitness testimony in international criminal tribunals.\textsuperscript{14} This Article dramatically expands that previous research and breaks new ground by providing quantitative assessments of two phenomena that are vitally important to accurate fact-finding. First, this Article presents a quantitative analysis of the testimonial deficiency in international criminal trials that is both most prevalent and that has the greatest potential to impair accurate fact-finding: inconsistencies between witnesses’ current testimony and their previous representations. Second, because testimonial deficiencies tell only part of the story of international criminal fact-finding, this Article quantitatively explores another, even more crucial part: the Trial Chambers’ treatment of witness testimony. I explore these issues through an empirical analysis of 342 prosecution witnesses who testified before the ICTR over the course of that Tribunal’s life.

Part II explains my research focus and my methodology. Part III details my findings on inconsistencies. To set the stage, I explain here what these inconsistencies are and how they arise. Before persons appear at an international criminal tribunal to testify for the prosecution, they tell their story at least once and often multiple times. At the very least, a prospective witness tells her story to a tribunal investigator who drafts a written statement, ostensibly containing the information that the prospective witness conveyed. Some prospective witnesses are interviewed multiple times and give multiple pre-trial statements. Additionally, some witnesses testify in multiple

\textsuperscript{11} See, e.g., Prosecutor v. Garda, ICC-02/05-02/09-243-Red, Decision on the Confirmation of Charges (Feb. 8, 2010); Prosecutor v. Mbarushimana, ICC-01/04-01/10-465-Red, Decision on the Confirmation of Charges (Dec. 16, 2011); Prosecutor v. Muthaura, ICC-01/09-02/11-382-Red, Decision on the Confirmation of Charges Pursuant to Art. 61(7)(a) and (b) of the Rome Statute, para. 425 (Jan. 23, 2012); Prosecutor v. Ruto, ICC-01/09-01/11-373, Decision on the Confirmation of Charges Pursuant to Art. 61(7)(a) and (b) of the Rome Statute, para. 293 (Jan. 23, 2012).

\textsuperscript{12} Defendants Thomas Lubanga, Germain Katanga, and Jean-Pierre Bemba, were convicted, see Lubanga Judgment, supra note 5, para. 1358; Prosecutor v. Katanga, ICC-01/04-01/07-3436-ENG, Judgment pursuant to Art. 74 of the Statute, 658–59 (Mar. 7, 2014) [hereinafter Katanga Judgment]; Prosecutor v. Bemba, ICC-01/05-01/08-3343, Judgment pursuant to Art. 74 of the Statute, 364 (Mar. 21, 2016), whereas Mathieu Ngudjolo was acquitted, see Prosecutor v. Ngudjolo, ICC-01/04-02/12-3-ENG, Judgment pursuant to Art. 74 of the Statute, 197 (Dec. 18, 2012) [hereinafter Ngudjolo Judgment].

\textsuperscript{13} See Ngudjolo Judgment, supra note 12, paras. 115–123; Dermot Groome, No Witness, No Case: An Assessment of the Conduct and Quality of ICC Investigations, 3 PENN. ST. J.L. & INT’L AFF. 1, 4, 19–21 (2014); see also Alex Whiting, Investigations and Institutional Imperatives at the International Criminal Court, in THE FIRST GLOBAL PROSECUTOR: PROMISE AND CONSTRAINTS 140 (Martha Minow et al. eds., 2015) (noting that in both judgments issued to date “the judges set aside considerable parts of the prosecution’s evidence, finding that it lacked sufficient credibility”).

\textsuperscript{14} See infra text accompanying notes 28 and 29.
cases about the same set of events, so, by the time these witnesses testify for
the prosecution, they have already testified under oath either before another
Trial Chamber or before another court entirely. Inconsistencies arise when a
witness’s testimony diverges from the representations that appear in the wit-
ness’s pre-trial statements or previous testimonies.

These sorts of inconsistencies are by no means confined to testimony
before the ICTR. My previous research found such inconsistencies to be a
prevalent feature of witness testimony in all of the international criminal
tribunals I studied, and they pertain to a whole range of topics relevant to
the disposition of the trial. Some inconsistencies center on such details as
dates, distances, duration, and numbers,15 whereas others concern central

15. See, e.g., Prosecutor v. Akayesu, Case No. ICTR-96-4-A, Judgment, para. 236 (June 1, 2001)
403, 412–13 (June 7, 2001); Prosecutor v. Gacumbis, Case No. ICTR-01-64-T, Judgment, para. 123
(June 17, 2004) [hereinafter Gacumbis Judgment]; Prosecutor v. Kajelijeli, Case No. ICTR-98-44A-T,
No. ICTR-95-54A-T, Judgment, paras. 327, 339 (Jan. 22, 2004); Prosecutor v. Karera, Case No. ICTR-
cutor v. Muvunyi, Case No. ICTR-95-1A-T, Judgment, paras. 269–270 (Apr. 28, 2005) [hereinafter
27, 2000) [hereinafter Musema Judgment]; Prosecutor v. Muvunyi, Case No. ICTR-00-55A-T, Judgment,
[hereinafter Paulino de Jesus Judgment]; Prosecutor v. Semanza, Case No. ICTR-97-20-T, Judgment,
para. 171 (May 15, 2003); Prosecutor v. Simba, Case No. ICTR-01-76-T, Judgment, paras. 167, 169,
382 (Dec. 13, 2005) [hereinafter Simba Judgment]; Transcript of Continued Trial at 56, 59–60, Prose-
cutor v. Bagosora (July 3, 2003) (ICTR-98-41-T); Transcript of Continued Trial at 9, 22, 30, Prosecutor
v. Bagosora (July 2, 2003) (ICTR-98-41-T) [hereinafter Military I Transcript, July 2, 2003); Transcript
of Continued Trial at 43, 44, 46, Prosecutor v. Gacumbis (Aug. 6, 2003) (ICTR-01-64-T); Transcript of
Kamuhanda Transcript, Sept. 19, 2001); Transcript of Continued Trial at 29–31, Prosecutor v. Kamu-
handa (Sept. 4, 2002) (ICTR-99-54A-T); Transcript of Continued Trial at 17–18, 20, Prosecutor
(Jan. 31, 2006) (ICTR-01-74-T); Transcript of Continued Trial at 43–45, Prosecutor v. Muvunyi
(Apr. 19, 2004) (ICTR-95-1B-T) [hereinafter Muvunyi Transcript, Apr. 19, 2004]; Transcript of Continued
Trial at 4, 18–19, Prosecutor v. Muvunyi (Mar. 31, 2004) (ICTR-95-1B-T) [hereinafter Muvunyi
Transcript, Mar. 31, 2004]; Transcript of Continued Trial at 12, 22–23, Prosecutor v. Ndindabahizi
(SCSL-04-16-T); Transcript of Trial at 96, Prosecutor v. Brima (July 26, 2005) (SCSL-04-16-T) [hereinafter
AFCR Transcript, July 26, 2005); Transcript of Trial at 82–83, 96, Prosecutor v. Brima (July 21, 2005)
(SCSL-04-16-T) [hereinafter AFCR Transcript, July 21, 2005); Transcript of Trial at 85, Prosecutor v.
Brima (July 12, 2005) (SCSL-04-16-T); Transcript of Trial at 151–52, Prosecutor v. Brima (July 11, 2005)
(SCSL-04-16-T) [hereinafter AFCR Transcript, July 11, 2005); Transcript of Trial at 122–23, Prosecutor
v. Brima (June 30, 2005) (SCSL-04-16-T); Transcript of Trial at 132–35, Prosecutor v. Brima (June 23, 2005)
(SCSL-04-16-T) [hereinafter AFCR Transcript, June 23, 2005); Transcript of Trial at 26–27, 47–48, Prosecutor
v. Brima (June 22, 2005) (SCSL-04-16-T); Transcript of Trial at 53–64, Prosecutor v. Brima (June 21, 2005)
(SCSL-04-16-T); Transcript of Trial at 21–25, Prosecutor v. Brima (Apr. 20, 2005) (SCSL-04-16-T); Trans-
script of Trial at 52–59, Prosecutor v. Brima (Apr. 19, 2005) (SCSL-04-16-T); Transcript of Trial at
Trial at 79, Prosecutor v. Brima (Apr. 8, 2005) (SCSL-04-16-T); Transcript of Trial at 24–27, Prose-
cutor v. Brima (Apr. 6, 2005) (SCSL-04-16-T); Transcript of Trial at 43, Prosecutor v. Norman (Mar. 8,
2005) (SCSL-04-14-T); Transcript of Trial 14–19, 24–28, Prosecutor v. Norman (Mar. 4, 2005) (SCSL-
aspects of the crime\textsuperscript{16} and/or the defendant’s involvement in the crime.\textsuperscript{17} Sometimes, a witness’s statement will incriminate the defendant when her testimony does not, and sometimes it is the witness’s testimony that incriminates when her statement does not.\textsuperscript{18} Finally, some witnesses testify about the


defendant’s key involvement in the crime, even though the witness’s pre-trial statement, or series of statements, fails even to mention the defendant.\(^19\)

For my past research, I did make some limited efforts to quantify inconsistencies to get a rough idea of the scope of the problem,\(^20\) but for this research, I dramatically expanded those quantification efforts by creating a large and rich dataset and by considering a host of explanatory factors. Part II conveys my findings in considerable depth, so it suffices to note here some of the myriad issues that I explore. My research reveals the percentage of

---


\(^20\) See FACT-FINDING WITHOUT FACTS, supra note 7, at 118–21.
prosecution witnesses who testified inconsistently with their previous statements/testimonies, and more importantly, the percentage whose inconsistencies are sufficiently worrisome that we would call them "serious," a term I will subsequently define. In addition, I considered a variety of explanatory variables such as the witness's gender, ethnicity, accomplice status, and imprisonment status. Thus, for instance, Part II presents tabular data showing whether male witnesses were more likely to testify seriously inconsistently than female witnesses; whether Hutu witnesses were more likely to testify seriously inconsistently than Tutsi witnesses; and whether witnesses who were imprisoned for genocide crimes were more likely to testify seriously inconsistently than witnesses who were not. In addition to witness characteristics, I considered various factors regarding the type, number, and timing of the witness's previous representations. Consequently, Part II details correlations between seriously inconsistent testimony and (1) the number of previous statements/testimonies a witness provided; (2) the judicial system in which those statements/testimonies were provided; and (3) the time that elapsed between the statements/testimonies and the later inconsistent testimony. Finally, Part II considers time trends and presents the results of several regressions that took into account a comprehensive set of explanatory variables. These regressions reveal which factors are statistically significant predictors of serious inconsistencies.

Testimonial deficiencies such as serious inconsistencies provide important information about the evidentiary foundations of international criminal judgments, but they tell only part of the story of fact-finding. The more important part examines the Trial Chambers' responses to testimonial deficiencies. Whatever the quality of the evidence a Trial Chamber receives, it will not affect the soundness of its judgments so long as the Trial Chamber accurately assesses the quality of the evidence and finds facts in accordance with that assessment. In other words, deficient testimony need not lead to deficient fact-finding. For purposes of my 2010 book, I took a summary look at the Trial Chambers' treatment of testimonial deficiencies, and I reported some troubling findings. For one thing, although some Trial Chambers mentioned some testimonial deficiencies, many did so only passingly, and some did not mention them at all. In addition, when Trial Chambers did mention testimonial deficiencies, they seemed almost reflexively inclined to attribute them to innocent causes that had no negative bearing on the witness's credibility. For instance, some Trial Chambers invoked the witness's educational or experiential deficiencies to explain their failure to provide relevant information or to answer whole ranges of probative ques-

21. See infra text at note 64.
22. A regression refers to the average relationship between a dependent variable that a researcher is trying to explain and one or more explanatory variables.
23. See FACT-FINDING WITHOUT FACTS, supra note 7, at 180–83.
As for serious inconsistencies in particular, I concluded that Trial Chambers often "explain these away as products of the passage of time, the frailty of memory, and errors introduced by investigators and interpreters. The Trial Chambers thus give the prosecution witnesses the benefit of the doubt, and they explain away problematic features of their testimony on the basis of innocent factors that are beyond the witnesses' control." In sum, although my prior research was not comprehensive or quantitative, it suggested that many Trial Chambers adopted a cavalier attitude toward testimonial deficiencies in general and serious inconsistencies in particular.

This Article quantifies the qualitative impressions just described. Specifically, in this Article, I have assessed each and every factual allegation brought by the prosecution in each of the nineteen cases studied. I have determined which witnesses' testimonies supported which allegations, and I have coded the Trial Chambers' conclusions regarding each witness's testimony. My findings are presented in Part III, and they paint a fascinating picture of the Trial Chambers' assessments of witness testimony and the factors that influence those assessments. Again, I have considered a variety of explanatory variables to assess whether gender, ethnicity, imprisonment status, accomplice status, or the presence of a serious inconsistency played a role in the Trial Chambers' treatment of witness testimony. In addition and most notably, I have considered the Trial Chambers' testimonial assessments over time and learned that the somewhat unflattering picture I presented six years ago has changed for the better. Part III provides the relevant details, but suffice it to say here that during the course of the ICTR's life, its Trial Chambers became increasingly skeptical of prosecution witness testimony and subjected that testimony to increasingly rigorous scrutiny.

Part IV seeks to explain the study's findings, many of which are surprising. The incidence of serious inconsistencies did not change as I expected, and it was not correlated with as many explanatory variables as I expected. The data do, however, reveal important clues about the most significant question surrounding serious inconsistencies: their cause. Obviously, multiple causes underlie a phenomenon as multi-faceted as serious inconsistencies; however, the data do suggest that some causes play a more prominent role than others. In particular, the data indicate that defense counsel may have had it right all along when they claimed that serious inconsistencies usually reflect false testimony. Turning to the big picture, the research, taken in its entirety, depicts a tribunal that confronted a relatively constant stream of evidentiary deficiencies but that altered—for the better—its responses to those deficiencies. The reasons for that evolution cannot be ascertained with certainty, but Part IV identifies both external circumstances that likely played a role as well as an internal maturation process that encompasses the

24. Id. at 190.
ICTR but that extends far beyond it. Indeed, the research presented here has particularly broad implications for the international criminal justice project as a whole, many of which are explored in this Article’s conclusion.

II: EXPLANATIONS AND METHODOLOGY

Section A explains why I chose the particular topics that form the basis of this study and how this research will enhance understanding of international criminal fact-finding. Section B details my methodology.

A. Explaining the Research Focus: A Spotlight on Serious Inconsistencies and Judicial Assessments of Witness Testimony

In past work, I identified a host of testimonial deficiencies that challenge international criminal fact-finding. Having decided in this piece to explore fact-finding through a rigorous empirical study, I necessarily had to narrow my focus. However, such a narrowing has the potential to distort, first because fact-finding is a holistic endeavor that cannot be reduced to a few isolated phenomena, and second because the phenomena that form the basis for the study, if not chosen correctly, can take on unjustified importance while at the same time inappropriately minimizing other, perhaps equally important factors. Because that potential for distortion exists, I will explain my research focus in some detail.

First, I chose to study serious inconsistencies because I consider them to be the most prevalent and pernicious testimonial deficiency challenging international criminal fact-finding. Why are serious inconsistencies so pernicious? Most obviously, it is because they call into question the accuracy of the testimony in which they appear. It goes without saying that witness testimony that sharply diverges from a witness’s previous representations is testimony that is less reliable and probably less likely to be accurate than testimony that does not so diverge. Inconsistencies, therefore, introduce considerable uncertainty into fact-finding.

To be sure, serious inconsistencies are not the only evidentiary phenomena that create uncertainty. For instance, my prior research revealed that international witnesses often do not know the answers to key questions that fact-finders need to ask in order to determine with any sort of certainty the who, what, where, and when details of the crimes in question. It also highlighted numerous instances in which witnesses appeared to know answers to relevant questions but were unable or unwilling to convey those answers in a

28. FACT-FINDING WITHOUT FACTS, supra note 7, at 21–44.
way that was comprehensible to their Western interlocutors. Serious inconsistencies, in my view, have a greater potential to impair accurate fact-finding than some of these other deficiencies, for reasons I explain below. Yet upon closer examination, it becomes clear that the primary impediment challenging accurate fact-finding at the international tribunals is not the existence of one or another individual testimonial deficiencies, but instead the very prevalence of witness testimony in international criminal trials. In short, witness testimony usually forms the exclusive basis for international criminal convictions, and that in itself is a problem.

It is a problem because eyewitness testimony is frequently unreliable, and indeed such testimony has been blamed for a large proportion of the wrongful convictions that have come to light in recent years. However problematic it is for the international tribunals to rely almost exclusively on witness testimony, the problem has had little remedy, at least historically, because little non-testimonial evidence of international crimes has typically

29. Id. at 44-62.
30. Id. at 11-14; see also THIERRY CRUVELLIER, COURT OF REMORSE: INSIDE THE INTERNATIONAL CRIMINAL TRIBUNAL FOR RWANDA 20 (Chari Voss trans., 2010). There are some exceptions to this rule. For instance, the Khmer Rouge, like the Nazis, documented many of their atrocities, so the Extraordinary Chambers in the Courts of Cambodia (ECCC) did receive large quantities of documentary evidence in Case 001. Prosecutor v. Kaing Guek Eav “Duch,” Case No. 001/18-07-2007/ECCC/TC, Judgment, para. 56 (July 26, 2010); HUMAN RIGHTS CTR., UNIV. OF CAL., BERKELEY, SCH. OF LAW, BEYOND REASONABLE DOUBT: USING SCIENTIFIC EVIDENCE TO ADVANCE PROSECUTIONS AT THE INTERNATIONAL CRIMINAL COURT 5 (2012) ("[T]he Court has relied heavily on documentary evidence, including lists of prisoners who were executed, photographs, and annotations written on 'confessions' of prisoners by their torturers."). Moreover, the ICTY did make more use of some non-testimonial evidence than other current international tribunals have. See, e.g., Prosecutor v. Prlić, Case No. IT-04-74-T, Judgment, para. 268 (May 29, 2013) (noting that the Chamber admitted a total of 9,756 items of documentary evidence); Prosecutor v. Popović, Case No. IT-05-88-T, Judgment, para. 260 n.835 (June 10, 2010); Prosecutor v. Kupreškić, Case No. IT-95-16-T, Judgment, paras. 184, 278, 395, 489(b) & nn.226, 309, 510, 507 (Jan. 14, 2000). In addition, prosecutors in the ICC case against Laurent Gbagbo have promised not only to introduce the testimony of 138 witnesses but also certain government documents. See Tom Maliti, Prosecutor: We Have Evidence to Prove Case against Gbagbo and Ble Goude!, INT’L JUST. MONITOR (Jan. 28, 2016), http://www.ijmonitor.org/2016/01/prosecutor-we-have-evidence-to-prove-case-against-gbagbo-and-ble-goude/. By and large, however, facts are proven at the international tribunals through witness testimony and fact witness testimony; in each of the SCSL’s first two trials, the prosecution presented only three expert witnesses out of seventy-five and fifty-nine prosecution witnesses, respectively. Prosecutor v. Fofana & Kondewa, Case No. SCSL-04-14-T, Judgment, Annex F, para. 21 (Aug. 2, 2007) [hereinafter CDF Judgment]; Prosecutor v. Brima et al., Case No. SCSL-04-16-T, Judgment, paras. 10, 149 (June 20, 2007). Some early ICTR cases similarly featured a few expert witnesses, but many of the more recent cases have featured only one or none at all. As for the ICC, the prosecution called only three expert witnesses out of thirty-six in Lubanga, see Lubanga Judgment, supra note 5, para. 11, and it called no expert witnesses in Katanga, see Katanga Judgment, supra note 12, para. 21.
31. FACT-FINDING WITHOUT FACTS, supra note 7, at 14-15. More recent research has only confirmed the problems associated with eye-witness testimony. See, e.g., BRANDON L. GARRETT, CONVICTING THE INNOCENT: WHERE CRIMINAL PROSECUTIONS GO WRONG 46-83 (2011); Deborah Davis & Elizabeth F. Loftus, The Dangers of Eyewitnesses for the Innocent: Learning from the Past and Projecting into the Age of Social Media, 46 NEW ENG. L. REV. 769, 769-74 (2012).
32. See GARRETT, supra note 31, at 8-9, 48 (finding that eyewitnesses misidentified 76 percent of the 250 exonerates in the author’s study); Brandon L. Garrett, Judging Innocence, 108 COLUM. L. REV. 55, 78-79 (2008); Eyewitness Misidentification, INNOCENCE PROJECT, http://www.innocenceproject.org/causes/eyewitness-misidentification/ (reporting that 72 percent of wrongful convictions are caused by eyewitness misidentifications).
been available to most international tribunals. Unlike Nazi war criminals who left carefully crafted, meticulous documentation of their atrocities, most modern day mass killers leave few written records, and because trials of international crimes often take place many years, if not decades, after the crimes occurred, most international tribunals likewise receive little forensic evidence. Furthermore, most modern atrocities occur in places that do not feature the widespread use of documentation or technology that can be so useful in proving a person’s whereabouts or other basic facts. Indeed, alibis are wildly popular defenses at many international tribunals probably because they can be plausibly claimed through a few corroborating witnesses. That is, whereas a defendant in a Western criminal trial who put forth an alibi defense would be expected to present receipts, ATM statements, or similar documentation to prove his presence at the claimed location, no such expectations exist with respect to a Rwandan defendant claiming an alibi because the Rwandan defendant truly might have spent considerable time in a location without generating documentary evidence to prove that he was there. Even ascertaining who is who can prove problematic at an interna-

33. In a forthcoming empirical study, I document the way in which criminal evidence is changing, particularly in developing societies and particularly with respect to non-testimonial evidence. In short, prosecutors of mass atrocities in developing nations have begun gaining access to greater quantities of non-testimonial evidence, and as the forthcoming piece explains, access to this evidence is poised to transform international criminal fact-finding.

34. See Prosecutor v. Kayishema, Case No. ICTR-95-1-T, Judgment, para. 65 (May 21, 1999) [hereinafter Kayishema Judgment]. As I explain in a forthcoming piece, Deconstructing the Epistemic Challenges to Mass Atrocity Prosecutions, the prevalence of non-testimonial evidence of international crimes is correlated with the development status of the location where the crimes took place. As a general matter, trials of international crimes occurring in developed nations feature more non-testimonial evidence than trials of international crimes occurring in developing nations. Because most recent international tribunal prosecutions of international crimes have centered on crimes occurring in developing nations, the bodies prosecuting these crimes have had little access to non-testimonial evidence.

35. ICTY Trial Chambers did receive non-trivial quantities of forensic evidence. It was unusual in that regard among international tribunals. Virtually the only forensic evidence submitted to ICTR Trial Chambers was introduced to prove that a genocide occurred: that is, it proved only that certain large-scale massacres did take place and that the victims of those massacres were Tutsi. See, e.g., Kayishema Judgment, supra note 34, paras. 325–26, 432. But see Prosecutor v. Ntagerera, Case No. ICTR-99-46-T, Judgment and Sentence, paras. 245, 252, 259, 260 (Feb. 25, 2004) (noting testimony by eyewitnesses and medical examiners that several exhumed bodies were identifiable as the remains of specific individuals). Alison Des Forges and Timothy Longman maintain that, “investigators made no systematic effort to gather documentary and forensic evidence linking alleged suspects to specific crimes.” Alison Des Forges & Timothy Longman, Legal Responses to Genocide in Rwanda, in MY NEIGHBOR, MY ENEMY: JUSTICE AND COMMUNITY IN THE AFTERMATH OF MASS ATROCITY 49, 53 (Eric Stover & Harvey M. Weinstein eds., 2004). The SCSL likewise received virtually no forensic evidence. But see Transcript of Open Session at 39–47, Prosecutor v. Norman (June 20, 2005) (SCSL-04-14-T) (testimony of forensic anthropologist William Haglund who, after examining the remains of four victims, determined that they had died from injuries that were consistent with their relatives’ descriptions of events).

36. See FACT-FINDING WITHOUT FACTS, supra note 7, at 162–65 (reporting that, as of 2010, more than 81 percent of ICTR defendants proffered alibis and that two of three SCSL cases then decided featured alibis).

37. In many countries like Rwanda, Sierra Leone, or the Democratic Republic of the Congo, commercial exchanges are carried out by means of cash or goods. Many transactions, particularly in rural areas, take the form of barter. See, e.g., Paul Cleary, It's the Economy Stupid, WALL ST. J. (May 23, 2007), http://www.wsj.com/articles/SB117987369444611381; Supporting Economic Growth in East Timor, THE ASIA
tional criminal trial because international witnesses frequently do not have birth certificates or other probative forms of official identification. 38

This lack of non-testimonial evidence renders accurate fact-finding more difficult, and it also increases the distortive potential of all the testimonial deficiencies I have discussed, including and perhaps especially serious inconsistencies. To be sure, fact-finding is a holistic process during which fact-finders take account of a host of relevant factors. The extent to which a witness’s testimony diverges from her previous representations is certainly one of those factors, but it is not the only one, and it may not even be a particularly important one in a trial that features a substantial quantity of non-testimonial evidence. For instance, although any criminal defense attorney would cross-examine a prosecution witness about inconsistencies between the witness’s testimony and her previous representations,39 the availability of non-testimonial evidence, such as documents, surveillance videos, wiretaps, and electronic data, provide defense counsel with other avenues to undermine the witness’s testimony. For example, it may not matter much that a witness’s testimony sharply diverges from her pre-trial statement if her testimony is strongly corroborated by a surveillance video, or—even more to the point—if the witness’s testimony is flatly contradicted by the video. In those instances, serious inconsistencies may remain a component of the credibility/reliability assessment that the fact-finder must conduct, but the other, more probative evidence will render that component an insignificant one.

38. For instance, in the Democratic Republic of the Congo, Central African Republic, and Sierra Leone—all locations forming the subject of international criminal trials—only 31 percent, 49 percent and 51 percent of births are registered, respectively. The registration rate for Rwandan births is 82 percent. See UNICEF, The State of the World’s Children 2016: A Fair Chance for Every Child 150–52 tbl.9 (2016).

39. See Kenneth J. Melilli, Examination of a Witness Based on a Prior Statement, 49 Am. Jur. Trials 501, § 2 (2013) (“There is perhaps nothing more potentially devastating on cross-examination than impeaching a witness with that witness’s own prior statement.”).
By contrast, in international tribunal proceedings—where witness testimony is often the only evidence presented\(^{40}\)—inconsistencies inevitably play a much more prominent role. Defense counsel seeking to undermine prosecution witness testimony but having no non-testimonial evidence by which to do so, frequently place inconsistencies at the center of their cross-examination.\(^{41}\) To be sure, defense counsel also advance their clients’ cases in other important ways. They present their own witnesses, who often contradict prosecution witness testimony, and they seek to undermine prosecution witness testimony in ways unrelated to inconsistencies. For instance, defense counsel expose prosecution witness testimony that seems improbable,\(^{42}\) they highlight incentives that might motivate witnesses to falsely inculpate the defendant,\(^{43}\) and they point out inconsistencies between witnesses.\(^{44}\) Despite the fact that international defense counsel unquestionably employ a multi-pronged approach when challenging prosecution witness testimony, they also unquestionably focus considerable attention—perhaps the lion’s share—on prosecution witness testimony that is inconsistent with previous

---


41. Even in the prosecution of domestic crimes, it has been recognized that “impeachment by means of prior inconsistent statement... is one of the most effective means of impeachment.” Michael H. Graham, *Impeachment of Witness—Prior Inconsistent Statements*, 21 AM. JUR. PROOF OF FACTS 2D 101, § 1 (2015); see also ROBERTO ARON & JONATHAN L. ROSNER, *HOW TO PREPARE WITNESSES FOR TRIAL* (1985).

42. See, e.g., Prosecutor v. Ndindabahizi, Case No. ICTR-2001-71-I, Judgment and Sentence, para. 191 (July 15, 2004) (noting that the “Defence challenges the plausibility and consistency of Witness CGH’s testimony that he was present in the room during the Accused’s visit to Bourgmestre Karara’s house,” questioning why the witness would risk being seen by the Accused, who knew him to be a Tutsi, when he had a “room that was reserved for him as a hiding place”); id. at paras. 195, 241; Prosecutor v. Sesay, Case No. SCSL-04-15-T, Sesay Defense Final Trial Brief, paras. 276, 1068 (Aug. 7, 2008) (hereinafter Sesay Defense Final Trial Brief).

43. For ICTR cases, see, e.g., Prosecutor v. Bizimungu et al., Case No. ICTR-99-50-T, Judgment and Sentence, para. 494, n.716 (Sept. 30, 2011) (hereinafter Bizimungu Judgment) (noting that defense counsel made allegations and presented witnesses to show that Witness GJQ had a pattern of lying in order to further his own self-interest and providing examples); Prosecutor v. Nакитиритамина, Case No. ICTR-96-17, Judgment and Sentence, para. 167 & n.211 (Feb. 21, 2003) (“[T]he Defence suggested that the witness was testifying against the Accused in the expectation that he would gain an early release from the Rwandan authorities.”). As for the SCSL, defense counsel in Prosecutor v. Brima [hereinafter AFRC case] alleged that insider witness Gibril Massaquoi had been informed by SCSL investigators that he was a suspect, and he agreed to testify in order to escape an indictment, see Kyra Sanin, U.C. Berkeley War Crimes Studies Center, Special Court Monitoring Program, Update No. 58, Oct. 10, 2005, at 4, and defense counsel in Prosecutor v. Sesay [hereinafter RUF case] made similar claims regarding insider witness John Tarnue, Transcript of Trial at 10, Prosecutor v. Sesay (Oct. 7, 2004) (SCSL-04-15-T).

statements/testimonies. Although I did not gather data on the proportion of cross-examination time that defense counsel devoted to probing inconsistencies, I can say with certainty that, in most cases, such probing occupied a substantial proportion of the cross-examination. Prosecutors likewise spend considerable time probing inconsistencies in the testimony of defense witnesses.

The parties focus so much attention on inconsistencies for three reasons. The first has already been discussed: the lack of non-testimonial evidence in international criminal trials leaves counsel with limited avenues for calling into doubt the credibility of witnesses and the reliability of their testimony. So, with few alternatives available, inconsistencies take center stage. Second, addressed more thoroughly in Part III, is the prevalence of inconsistencies. That is, counsel focus on inconsistencies because there are a lot of inconsistencies on which to focus—both at the ICTR and elsewhere. As Part III reveals, 67 percent of prosecution witnesses in the nineteen ICTR cases I studied presented testimony that was in some way inconsistent with their previous representations, and nearly 50 percent testified in a way that was seriously inconsistent. Third and finally, ICTR counsel in particular focus on inconsistencies because a great deal of ICTR testimony features mistakes or lies, and scrutinizing inconsistencies may help to separate the accurate from the inaccurate.

It is apparent that a great deal of ICTR testimony features mistakes or lies because ICTR trials are filled with witnesses who blatantly contradict one another. My previous research has shown that more than 90 percent of ICTR cases featured at least one blatant contradiction between witnesses for the defense and witnesses for the prosecution. Many cases featured far more than one, and most multiple-defendant cases featured at least one contradiction relating to each defendant. I considered witness testimony to be blatantly contradictory only where the testimony of one witness was diametrically opposed to the testimony of another, such that both witnesses’

---
45. See, e.g., Bagilishema Judgment, supra note 15, para. 549 (“In the absence of details, the Chamber has looked into the witness’s previous written statements.”).
46. At the SCSL, for instance, 54 percent of witnesses in the AFRC case testified seriously inconsistently; so too did 53 percent of witnesses in the RUF case and 35 percent in Prosecutor v. Norman [hereinafter CDF case]. Fact-Finding Without Facts, supra note 7, at 118-19.
47. Some empirical research shows greater levels of inconsistencies among liars. One particularly notable study showed that liars were “significantly more likely than truth tellers to add details to later statements that they had not mentioned in earlier statements.” Anneliese Vredenvelde, et al., The Inconsistent Suspect: A Systematic Review of Different Types of Consistency in Truth Tellers and Liars, in Investigative Interviewing 183, 189-93 (Ray Bull, ed., 2014). However, other studies show liars to be equally or even more consistent than truth tellers. See id. at 193-94.
50. Id. at 394.
allegations could not possibly be true. When witnesses contradict one another in this way, it becomes apparent that one or the other witness was necessarily testifying inaccurately, either on purpose or by mistake. As noted, ICTR cases virtually never featured any non-testimonial evidence to assist in determining which witness’s testimony was inaccurate, so the parties naturally looked to inconsistencies. That is, when a prosecution witness testifies that the defendant led a massacre, and defense witnesses claim that the defendant was with them hundreds of miles from the massacre site, and no other evidence of the defendant’s whereabouts during the massacre is available, then key differences between the prosecution witness’s testimony and her previous representations understandably take on crucial significance.

In sum, inconsistencies play a central role in this study because they play a central role in international criminal trials. Inconsistencies stand as the most high-profile testimonial deficiency confronting international tribunals and the deficiency that provides particularly useful information when witnesses testify contradictorily. To understand international criminal trials, then, a deeper understanding of the inconsistencies that feature so prominently in these trials must be gained.

My focus on the Trial Chambers’ assessments of witness testimony needs less explanation. The accuracy of a court’s factual findings is of central concern to all court watchers and participants. Although it is impossible to determine whether a court’s findings are accurate when the facts are contested, it is possible to evaluate the court’s fact-finding methodology. What are the characteristics of the witnesses whom the Trial Chambers credit? How often and for what reasons do the Trial Chambers decline to rely on prosecution witness testimony? The answers to these and similar questions provide important insights into the court’s fact-finding methodology, and that methodology helps us to assess the likely accuracy of the court’s factual findings. Scholars of any criminal justice system desire the answers to those questions, but often they cannot get them. American scholars, for instance, can learn little about the credibility and reliability assessments made during a criminal trial because American criminal trials end with a jury verdict of guilty or not guilty, and no further information about the jury’s assessment of the evidence is available. International tribunals, by contrast, issue extraordinarily long judgments that typically detail all of the evidence presented and the Trial Chamber’s assessment of that evidence. Although many criticize the long length of tribunal judgments,51 one benefit of such careful detailing of the evidence is that it provides scholars with a trove of data that can help us to better understand the way in which international tribunals carried out what may be their most important function: fact-finding.

Finally, this section concludes by explaining why the Article centers on ICTR cases, rather than the cases of another international criminal tribunal. In order to have confidence in these findings, a reasonably large number of cases is needed. That need immediately eliminates the ICC, the SCSL, and the Special Tribunal for Lebanon from consideration because they have decided, at most, only a few cases each. This leaves only the ICTR, the ICTY, and the Special Panels for Serious Crimes in East Timor (Special Panels). Although I have previously researched fact-finding at the Special Panels, and found that research to reveal many important insights; for this empirical study, Special Panels cases would not have been appropriate. Special Panels transcripts are not readily available, and Special Panels judgments, if they exist at all, are exceedingly short and undetailed. ICTY cases also would not have provided an optimal dataset. For one thing, ICTY witnesses are something of an outlier in international criminal trials because, on average, they are markedly better educated than witnesses appearing before other international tribunals, and they are more likely to have had life experiences that enable them to answer the kinds of questions typically posed in a criminal trial. For that reason, ICTY trials featured fewer testimonial deficiencies. Moreover, ICTY Trial Chambers received more non-testimonial evidence than most other international tribunals, and this evidence served both to reduce the importance of witness testimony at the ICTY and concomitantly the deficiencies common to such testimony. Thus, I could have studied serious inconsistencies in ICTY witness testimony because there were in fact at least some ICTY witness testimony that featured such inconsistencies. However, the far lower incidence of inconsistencies at the ICTY and their reduced influence would likely have produced a distorted picture that would not have generalized to fact-finding in other, current tribunals, which unfortunately feature more problematic witness testimony and less non-testimonial evidence. The ICTR, by contrast, decided a sufficient number of cases that featured the kind of evidence that also appears in current tribunals, such as the ICC. The size of the dataset that I was able to create gives me confidence in my findings, and the evidentiary profile of the ICTR cases gives me confidence in their generalizability across many international tribunals.

B. Methodology

1. The Cases

As noted in the Introduction, my dataset comprises nineteen single-defendant cases from the ICTR. Table 1 below shows the cases in chronologi-
cal order along with the starting date of each trial, the ending date, and the date on which the Trial Chamber issued its judgment in that case.

**Table 1: Starting, Ending, and Judgments Dates for Each Trial in Chronological Order**

<table>
<thead>
<tr>
<th>Trial</th>
<th>Starting Date of the Trial</th>
<th>Ending Date of the Trial</th>
<th>Date of the Trial Chamber's Judgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rutaganda</td>
<td>Mar. 18, 1997</td>
<td>June 17, 1999</td>
<td>Dec. 6, 1999</td>
</tr>
<tr>
<td>Gacumbitsi</td>
<td>July 28, 2003</td>
<td>Mar. 1, 2004</td>
<td>June 17, 2004</td>
</tr>
<tr>
<td>Muhimana</td>
<td>Mar. 29, 2004</td>
<td>Jan. 20, 2005</td>
<td>Apr. 28, 2005</td>
</tr>
<tr>
<td>Setako</td>
<td>Aug. 25, 2008</td>
<td>Nov. 6, 2009</td>
<td>Feb. 25, 2010</td>
</tr>
<tr>
<td>Hategekimana</td>
<td>Mar. 16, 2009</td>
<td>Apr. 28, 2010</td>
<td>Dec. 6, 2010</td>
</tr>
<tr>
<td>Ntawukulilyayo</td>
<td>May 6, 2009</td>
<td>June 14, 2010</td>
<td>Aug. 3, 2010</td>
</tr>
<tr>
<td>Kanyarukiga</td>
<td>Aug. 31, 2009</td>
<td>May 24, 2010</td>
<td>Nov. 1, 2010</td>
</tr>
</tbody>
</table>

To facilitate the exploration of any relevant time trends, the dataset includes cases from the beginning, middle, and end of the Tribunal's life. In addition, in order to ensure a representative sample, the dataset includes cases featuring defendants who had high-level, mid-level, and low-level positions and defendants who participated in the genocide in a variety of ways and across a variety of regions in Rwanda.54 Seventy-four-year-old55 busi-

---

54. The nineteen cases in my dataset featured crimes in all but two of Rwanda's then-twelve prefectures (Gikongoro and Umurara). Crimes in about half of the prefectures (Butare, Gitarama, Kibungo, Kibuye, Kigali-Rural, Kigali-Ville, and Ruhengeri) were the subject of two or more cases.

55. Specifically, Munyakazi was seventy-four years old at the time his judgment was pronounced in 2010. Prosecutor v. Munyakazi, Case No. ICTR-97-36-A-T, Judgment and Sentence, para. 29 (July 5, 2010) [hereinafter Munyakazi Judgment].
nessman and farmer, Yussuf Munyakazi, for instance, was sufficiently low-level that ICTR prosecutors sought (unsuccessfully) to refer his case to Rwanda before eventually trying him at the ICTR. However, the dataset also features such high-ranking defendants as Eliezer Niyitegeka and Augustin Ngirabatware, who were Ministers in the Interim Government of Rwanda during the genocide. The remaining defendants held a host of other positions, some in the government, some in the military, and one in the Catholic Church.

The cases were decided by a total of twenty-eight judges. More particularly, all of the cases were tried before a panel of three judges, and sixteen of the judges participated in more than one trial in the dataset. Each of the defendants in the dataset was convicted of at least one of the charges leveled against him. Cases in which defendants were acquitted of all charges were not included. This dataset involves only prosecution witness testimony, and prosecution witness testimony in trials ending in acquittals likely feature


62. The judges who participated in more than one trial were Judges Akay, Arrey, Bossa, Egorov, Hikmet, Khan, Masanche, Mose, Muthoga, Park, Pillay, Rajohnson, Ramaroson, Sekule, Tuzmukhamedov, and Vaz.
greater deficiencies than prosecution witness testimony in ICTR trials as a whole. Inclusion of these cases, therefore, could have biased the sample.

2. The Witnesses

The nineteen cases in the dataset featured 342 prosecution fact witnesses. The gender, ethnicity, and other relevant characteristics of the 342 witnesses are shown in Tables 2 through 5.

**Table 2: Gender Composition of Witnesses in the Dataset**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>255</td>
<td>74.56%</td>
</tr>
<tr>
<td>Female</td>
<td>86</td>
<td>25.15%</td>
</tr>
<tr>
<td>Unable to Ascertain</td>
<td>1</td>
<td>0.29%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>342</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Table 3: Ethnic Composition of Witnesses in the Dataset**

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hutu</td>
<td>133</td>
<td>38.89%</td>
</tr>
<tr>
<td>Tutsi</td>
<td>192</td>
<td>56.14%</td>
</tr>
<tr>
<td>Neither</td>
<td>11</td>
<td>3.22%</td>
</tr>
<tr>
<td>Unable to Ascertain</td>
<td>6</td>
<td>1.75%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>342</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Table 4: Accomplice Status of Witnesses in the Dataset**

<table>
<thead>
<tr>
<th>Accomplice Status</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accomplices of the Defendant</td>
<td>67</td>
<td>19.59%</td>
</tr>
<tr>
<td>Not Accomplices of the Defendant</td>
<td>266</td>
<td>77.78%</td>
</tr>
<tr>
<td>Unable to Ascertain</td>
<td>9</td>
<td>2.63%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>342</td>
<td>100%</td>
</tr>
</tbody>
</table>
The 342 witnesses in the dataset comprise all of the prosecution fact witnesses who testified in the trials listed in Table 1. The average number of prosecution fact witnesses per case was eighteen. Kanyarukiga had the lowest number, at ten, and Nizeyimana the highest, at forty-one. The vast majority of witnesses in the dataset, 97 percent, were Rwandan.

Why did I choose to study prosecution witnesses instead of defense witnesses or instead of all of the witnesses in a given case? The decision to focus on only one side’s witnesses was driven by both practical and substantive concerns. As a practical matter, I had the time to assess the testimony of only a certain number of witnesses. If I had included in the database all of the witnesses in each of the cases, then I would have had to reduce the number of cases that I evaluated. I found that prospect undesirable because I considered it important to include cases that featured a variety of different kinds of defendants and cases that were tried during different periods in the Tribunal’s life. Thus, I believed that including more cases would produce more certain results, even though that meant including a smaller number of witnesses per case. Having made the decision to include witnesses from only one side, I selected prosecution witnesses for two reasons. First, if we assume, as most commentators do, that wrongful convictions are a greater injustice than wrongful acquittals, then we have more reason to be concerned about problematic prosecution testimony than problematic defense testimony. Moreover, although I have not systematically analyzed defense witness testimony, my non-quantitative sense is that it features even more serious inconsistencies than prosecution witness testimony. If that is correct, then an exclusive focus on defense witness testimony would overstate the problematic features of international tribunal testimony. As it stands, my

---

63. 4 WILLIAM BLACKSTONE, COMMENTARIES ON THE LAWS OF ENGLAND 352 (1769); William O. Douglas, Foreword to JEROME FRANK & BARBARA FRANK, NOT GUILTY 11 (1957). As Thomas Starkie put it, the “maxim of the law is . . . that it is better that ninety-nine . . . offenders shall escape than that one innocent man be condemned.” 9 JOHN WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 2497, 1650–51 (Chadbourn rev. 1981).
3. The Inconsistencies

a. Defining Inconsistencies

I defined an inconsistency as witness testimony that was inconsistent with the witness’s pre-trial statements or previous testimony in other cases. International tribunal testimony features numerous inconsistencies, but a non-trivial proportion are minor or concern minor details in the trial. Thus, although I gathered some data on all inconsistencies, I focused my research on “serious” inconsistencies.

This study considers an “inconsistency or omission to be serious either if it pertained to a key issue in the trial or if it pertained to the kind of fact that one is unlikely to forget.” The former category included inconsistencies or omissions “that directly related to the defendant’s actions or overall liability or to the witness’s credibility or the weight the Trial Chamber should place on the witness’s testimony.” So, for instance, if a witness’s statement described the defendant as engaging in a different sort of criminal behavior than that to which the witness testified, I considered that a serious inconsistency. One example would be a witness who testified that the defendant personally killed the victims when the statement reported that the defendant was merely present during the killings. I likewise deemed an inconsistency serious if it called into question the witness’s ability to observe the events she described. So, if the witness’s pre-trial statement reported that she was 100 meters from the crime site, but she testified that she was five meters away, I considered that a serious inconsistency. Finally, I considered it a serious inconsistency when a witness failed to mention in his previous statements/testimonies a fact that was central to his testimony. The most common example of this phenomenon occurred when a witness provided detailed statements that contained no reference to the defendant’s participation in the relevant crime, yet in later statements or sometimes only in

64. FACT-FINDING WITHOUT FACTS, supra note 7, at 121 (defining “serious inconsistency” in the same way).

65. Id.


testimony, the witness claimed that the defendant had played an integral role in the crime.\textsuperscript{68}

The second category of inconsistencies that I deemed to be serious included those pertaining to the kind of facts that the witness was "unlikely to forget" even if such facts were not of crucial significance to the resolution of the case. So, for instance, I considered it serious "if a defendant testified inconsistently about where he hid because one would expect that he'd remember whether he hid in the parish church, say, or in the bushes behind his house. Similarly, I considered it a serious inconsistency if a witness testified that she hid with one of her children when her statement reported that she hid with three of her children."\textsuperscript{69}

b. Methodology for Identifying Inconsistencies

The best method for identifying inconsistencies would be to read each witness's statements/testimonies and then compare the allegations contained therein with the witness's ICTR testimony. Unfortunately, the ICTR, like other international tribunals, places exhibits such as pre-trial statements and testimony before Rwandan courts under seal, so they are not publicly available. Consequently, those seeking to identify inconsistencies at the international tribunals must rely on discussions of those inconsistencies that occur during trial or descriptions of the inconsistencies that appear in the Trial Chambers' judgments. This method of locating inconsistencies is apt to understate them for two reasons. First, as a general rule, inconsistencies in the


\textsuperscript{69} \textbf{FACT-FINDING WITHOUT FACTS,} supra note 7, at 122. For a sampling of these sorts of inconsistencies, see, e.g., Gacumbitsi Judgment, supra note 15, at para. 192; Gatete Judgment, supra note 58, at paras. 96–97; Transcript of Continued Trial at 14–17, Prosecutor v. Gatete (Nov. 3, 2009) (ICTR-00-61-T); Transcript of Continued Trial at 53, Prosecutor v. Gatete (Nov. 5, 2009) (ICTR-00-61-T).
testimony of prosecution witnesses are mentioned only by defense counsel, and defense counsel almost certainly do not mention all of them. Second, not all testimony is public, so inconsistencies that are discussed in closed testimony and that are not later described in the Trial Chambers' judgment will not become known. As I will discuss below, I cannot remedy the first source of potential understatement, but I have sought to remedy the second.

Relying on defense counsel to highlight serious inconsistencies is sub-optimal for two reasons. First, the skill, diligence, and strategic intuitions that defense counsel bring to their representation vary considerably from attorney to attorney.70 So, although the transcripts show that most defense counsel vigorously question prosecution witnesses on perceived inconsistencies, in at least one case in my dataset, defense counsel seemed to eschew this cross-examination technique entirely.71 Second, ICTR trials are largely adversarial, so even when defense counsel are generally inclined to point out inconsistencies, they probably would mention only those whose identification would advance their clients' interests. To be sure, pointing out any inconsistency can serve to undermine a prosecution witness's credibility and reliability, so it may be that defense counsel bring to light most of the inconsistencies that they discover. However, it is reasonable to assume that defense counsel become aware of some inconsistencies that they choose not to point out. For instance, although we can be virtually certain that defense counsel will (enthusiastically) point out an inconsistency in which the witness's testimony implicates the defendant in more criminal activity than does the pre-trial statement, we cannot be so sure about the reverse. When a witness's testimony is less inculpatory than his pre-trial statement, defense counsel may see good reason to refrain from mentioning the more inculpatory pre-trial statement. Thus, because I learned of inconsistencies only when defense counsel referred to them, I have probably under-counted them to some degree. At the same time, any under-counting may have been ameliorated by the fact that defense counsel have an incentive to exaggerate inconsistencies; thus, they may describe allegations as inconsistent when they are not in fact. Or, more plausibly, they might describe actual inconsistencies in a way that overstates the importance or the degree of the divergence. Cognizant of this possibility, I carefully considered the witnesses' responses to defense counsel allegations and any subsequent re-direct testimony that the prosecution elicited. Finally, the Trial Chambers mentioned a

70. As Cruvellier put it, "[t]he defense lawyers at the ICTR would always be as heterogeneous as they were individual... Up to the very end, there would always be a mix of those who were greedy, and those who had integrity, the mediocre and the inspired, the nitpickers and the parsimonious, the pompous and the first-rate cross-examiners, the brilliant litigants and the boring..." Cruvellier, supra note 30, at 36.

71. In Munyakazi, defense counsel failed even to mention the statements of at least 70 percent of the prosecution witnesses. See generally Munyakazi Judgment, supra note 55.
substantial proportion of serious inconsistencies,72 so I cross-checked my assessment of the inconsistencies with that of the Trial Chambers.

The second reason that I likely under-counted inconsistencies stems from the fact that some transcripts are wholly unavailable and some testimony appearing in available transcripts is held in camera, or "closed session", so it does not appear in the publicly available transcripts.73 Testimony that is not publicly available may feature serious inconsistencies, but we will not know unless the Trial Chamber mentions them in its judgment. In order to take account of serious inconsistencies that were "hidden" in this way, I used the following formula (which extrapolates from the publicly available information that we do have) to add an estimated percentage of serious inconsistencies.

\[(P + .5Y&N)C,\]

where

- \(P\) is the percentage of witnesses in the case in question whose serious inconsistencies the Trial Chamber did not mention
- \(Y&N\) is the percentage of witnesses in that case for whom the Trial Chamber mentioned some of their serious inconsistencies but not others
- \(C\) is the percentage of cross-examination testimony held in camera in that case

I applied the formula on a case-by-case basis, and I added the result reached by the formula to the number of witnesses in that case who testified seriously inconsistently.

As the formula shows, the primary factors in my calculation are the percentage of closed cross-examination testimony in a given case and the percentage of serious inconsistencies that the Trial Chamber in that case failed to mention in its judgment. The formula includes only cross-examination testimony held in camera because defense counsel virtually always raised inconsistencies during cross-examination. Thus, in camera direct testimony was unlikely to have featured serious inconsistencies. The formula assumes that the larger the proportion of cross-examination held in camera, the more likely that serious inconsistencies have gone unidentified. However, Trial Chambers do mention some serious inconsistencies in their judgments, even when the inconsistencies are discussed only in in camera testimony, so the formula also takes account of the Trial Chambers' willingness to mention serious inconsistencies. In particular, the formula assumes that the larger the proportion of witnesses with serious inconsistencies whom the Trial Chambers mentioned, the less likely that a serious inconsistency discussed in closed testimony went unidentified. As it happens, Trial Chambers varied considerably in their willingness to mention serious inconsistencies, which is

72. See infra text accompanying notes 127 and 128.
73. For the argument that excessive use of in camera testimony denies defendants the right to a public trial, see Transcript of Trial Hearing at 27, Prosecutor v. Ntaganda (Sept. 3, 2015) (ICC-01/04-02/06-T-24-ENG).
why I calculated the formula on a case-by-case basis. Finally, I added half
the percentage of witnesses labeled Y&N in that case, estimating that Trial
Chambers mentioned half of the serious inconsistencies of those labeled
Y&N. Because few witnesses were labeled Y&N, and because I believe the
rate of reference is about 50 percent for such witnesses, I did not calculate
the percentage on a per-witness basis.

The following example will illustrate the operation of the formula. As-
sume that through a review of the publicly available testimony, I deter-
mind that ten out of twenty-five prosecution witnesses testified seriously
inconsistently in a particular case. Assume also that, for three witnesses in
this case, some or all of their testimony occurred in camera. In particular,
assume that all of Witness A's testimony was held in camera, one-half of
Witness B's testimony, and one-tenth of Witness C's testimony was held in
camera. Assume also that the judgment in this case failed to reference the
serious inconsistences of 25 percent of the witnesses who we know (from
reading the publicly available transcripts) in fact testified seriously inconsist-
tently. Assume finally that for 4 percent of the witnesses in this case who
testified seriously inconsistently, the Trial Chamber referenced some but not
all of their serious inconsistencies. Assuming these facts, we would first ag-
gregate the proportion of cross-examination testimony for each witness that
was held in camera. The aggregation would look like this:

\[
1 \text{(Witness A)} + .5 \text{(Witness B)} + .1 \text{(Witness C)}
\]

Our sum is 1.6. We would also aggregate the percentage of witnesses whose
serious inconsistences were not mentioned (which is 25 percent) and one
half of the percentage of Y&N witnesses (which is 2 percent) to reach 27
percent. Then we would multiply 1.6 by 27 percent to reach .432 and add
.432 to the number of witnesses who we know testified seriously inconsist-
tently in that case—here, ten. Adding .432 to ten leads to an adjusted esti-
mate of 10.43 witnesses who testified seriously inconsistently. Finally, on
these facts, the percentage of witnesses who testified seriously inconsistently
would increase from 40 percent (10/25) to 41.7 percent (10.43/25). I label
the former figure the "understated percentage of serious inconsistencies"
and the latter figure the "adjusted percentage of serious inconsistencies."

I used the adjusted percentage of serious inconsistences for only two cal-
culations in my dataset. First, in reporting the percentage of witnesses who
testified seriously inconsistently, I presented both the understated and the
adjusted percentage of serious inconsistences. Second, Graph I shows the
percentage of witnesses who testified seriously inconsistently over time, and
here, I also used the adjusted percentage of serious inconsistences. For all
other calculations, I used the understated percentage of serious inconsist-
encies. I did so primarily because virtually none of my other findings would
have changed had I used the adjusted percentage of serious inconsistences.
Most of my calculations relating to serious inconsistencies compared their incidence in different populations of witnesses. So, for instance, I compared the percentage of female and male witnesses who testified seriously inconsistently and the percentage of Hutu and Tutsi witnesses who did so. Adding an estimated percentage to account for the serious inconsistencies that likely appeared in unavailable transcripts obviously increases the overall percentage of witnesses who testified seriously inconsistently, but my calculations showed that that increase is distributed close to evenly across the different sub-populations. Even though the differences are not substantial, I would have preferred to run regressions on the adjusted data, but because the regressions require categorical data, not continuous data, I had to use the understated percentage of serious inconsistencies.

4. Details Concerning Statements/Testimonies

In order to determine their influence on the incidence of serious inconsistencies, I gathered data on the number and types of statements/testimonies that witnesses made prior to testifying in the subject trial. Because pre-trial statements are filed under seal, I was able to learn about them only by reading the relevant judgments and transcripts. Unfortunately, this method almost certainly understated the number of statements/testimonies made by witnesses because defense counsel typically mentioned them only when cross-examining the witness about a perceived inconsistency between a particular statement and witness’s current testimony. Thus, I had no way to identify statements that defense counsel had no reason to mention. Based on the representations of a defense counsel whom I interviewed, I did, however, assume that each witness made at least one pre-trial statement for the ICTR.

I categorized the witnesses’ previous statements/testimonies into five groups: (1) pre-trial statements taken by ICTR investigators; (2) testimony in previous ICTR cases; (3) statements and testimony taken in genocide trials in Rwandan courts; (4) statements and testimony taken in genocide trials in foreign courts; and (5) statements and testimony taken in gacaca

---

74. For example, I found the largest difference between the understated percentage of serious inconsistencies and the adjusted percentage in my comparison on the basis of gender. In particular, the difference between the understated percentage of serious inconsistencies and the adjusted percentage was .18 for men and .27 for women. The differences for the other subpopulations—comprising the two ethnicities, the accomplices/non-accomplices, and the imprisoned/non-imprisoned—were less than half the difference for the two genders.

75. These regressions were trying to determine the effect of various factors, such as gender or ethnicity, on the probability of a witness’s testimony containing a serious inconsistency. Thus each observation involved one witness, and that witness either did or did not have a serious inconsistency. To make such calculations, the dependent variable must take on the value one for the existence of a serious inconsistency and zero if not. This approach does not allow for values of the dependent variables that are fractions between zero and one.

76. Skype Interview with Peter Robinson, Defense Counsel (June 16, 2015).
proceedings in Rwanda. Ideally, I would have subdivided categories (3), (4), and (5) and placed statements and testimony in each of these criminal justice systems into separate categories of their own. I was unable to do so, however, because the transcripts and judgments often did not provide sufficient detail. In a typical cross-examination, for instance, defense counsel might make reference to some allegations the witness previously made during gacaca, but the discussion would be unclear as to whether the allegations were made during a gacaca trial or in some sort of pre-trial statement. I do not consider my inability to refine the data in this way to be problematic, however, because my primary goal in considering this set of data was to ascertain whether documents generated by some criminal justice systems were more likely to give rise to inconsistencies than documents generated in other criminal justice systems. The data answer that question, as I will describe in Part III.

5. The Trial Chambers’ Assessments of Prosecution Witness Testimony

In considering the Trial Chambers’ assessments of witness testimony, I focused on their findings of witness credibility and their willingness to rely on witness testimony. Although many scholars (and some judges) devote considerable attention to probing the nuances of the distinction between “credibility” and “reliability” and the relationship between those concepts and the admissibility and weight of evidence, I use the following basic definitions. “Credibility” is typically equated with truthfulness, such that a witness who is testifying honestly can be deemed “credible.” “Reliability,” by contrast, can include credibility, but also encompasses the witness’s ability to observe the events about which he or she is testifying. Sometimes international tribunals appear to confuse credibility with reliability, but for purposes of this Article, I took the Trial Chambers at their word. Thus, I

77. Rwandan courts were unable to handle the prosecution of all of the offenders arrested following the genocide, so the Rwandan government adapted an indigenous dispute resolution process known as gacaca into a method for prosecuting genocide. PHIL CLARK, THE GACACA COURTS, POST-GENOCIDE JUSTICE AND RECONCILIATION IN RWANDA 55–63 (2010).

78. See, e.g., Prosecutor v. Kunare et al., Case No. IT-96-23 & 23/1, Decision on Motion for Acquittal, para. 7 (July 3, 2000); MARK KLAMBERG, EVIDENCE IN INTERNATIONAL CRIMINAL TRIALS: CONFRONTING LEGAL GAPS AND THE RECONSTRUCTION OF DISPUTED EVENTS 351–57 (2013); Mark Klamberg, General Requirements for the Admission of Evidence, in INTERNATIONAL CRIMINAL PROCEDURE: PRINCIPLES AND RULES 1016, 1025–29 (Görän Sluiter et al. eds., 2013); RICHARD MAY & MARIEKE WIERDA, INTERNATIONAL CRIMINAL EVIDENCE 107–11 (2002).

79. See, e.g., Prosecutor v. Brcanin, Case No. IT-99-36-T, Judgment, para. 25 (Sept. 1, 2004); CDF Judgment, supra note 30, at para. 257; Prosecutor v. Sesay et al., Case No. SCSL-04-15-T, Judgment, para. 487 (Mar. 2, 2009); KLAMBERG, EVIDENCE IN INTERNATIONAL CRIMINAL TRIALS, supra note 78, at 174–77; Klamberg, General Requirements for the Admission of Evidence, in INTERNATIONAL CRIMINAL PROCEDURE, supra note 78, at 1025; Groome, supra note 13, at 19 n.49.

80. See, e.g., Seromba Judgment, supra note 60, at para. 65 (finding that YAT’s testimony “cannot be deemed credible” because “the information which was disclosed to him [was] not supported by any other evidence”). Such confusion is not confined to ICTR Trial Chambers. See, e.g., Prosecutor v. Naletelić & Martinović, Case No. IT-98-34-A, Judgment, para. 402 (May 3, 2006); KLAMBERG, EVIDENCE IN INTERNATIONAL CRIMINAL TRIALS, supra note 78, at 174 (providing additional examples); Klamberg, Gen-
classified a witness as credible if the Trial Chamber deemed the witness's testimony to be "credible," even if I suspected—based on the Trial Chamber's additional commentary—that it was, or should have been, referring to reliability.

For purposes of this study, I characterized the Trial Chamber as finding a witness credible or not credible when the Trial Chamber either expressly stated that it found the witness to be credible or not credible, or when it was absolutely clear from the context and Trial Chambers' other findings about the witness that it considered the witness credible or not credible. For nearly 6 percent of witnesses, I categorized the Trial Chambers' findings as "yes and no," because the Trial Chamber expressly found the witness credible in some respects or for some purposes but not credible in other respects or for other purposes. For instance, the Renzaho Trial Chamber generally treated the testimony of witness ALG with caution, finding that "his evidence may have been influenced by a wish to positively affect the proceedings against him in Rwanda." In part for that reason, no doubt, the Trial Chamber did not find witness ALG credible with respect to his allegations about the dismissal of Conseiller Célestin Sezibera. Yet, it did find him to be "consistent and credible" when it came to his testimony regarding Renzaho's meetings in late February and early March. Finally, I was forced to withhold a credibility assessment for 16 percent of witnesses because the Trial Chambers simply did not provide sufficient information for me to determine whether they found these witnesses credible or not.

Classifying the Trial Chambers' reliance on testimony was more complicated. First, for each witness, I identified all of the allegations that the prosecution sought to prove through a particular witness's testimony. Thus, if the witness testified that the defendant spoke at a rally on April 10th, and delivered weapons on April 11th, and participated in a massacre on April 12th, then I listed each of these three allegations, and I determined whether the Trial Chamber relied on the witness's testimony to prove each of the three allegations. Each allegation, therefore, generated a "yes" or "no" entry. For most allegations, the classification was straightforward because the Trial Chamber clearly relied on the witness's testimony to prove each of the three allegations. Each allegation, therefore, generated a "true" "yes" or "no" entry. For most allegations, the classification was straightforward because the Trial Chamber clearly relied on the witness's testimony in finding the allegation to be proved or rejected the witness's testimony in finding the allegation not proved. These allegations generated a "true" "yes" or "no" entry. In addition, however, some allegations generated a "yes"—because the Trial Chamber relied on the witness's testimony to prove the allegation—or a

82. Renzaho Judgment, supra note 59, at para. 113 n.137.
83. See id. at paras. 494, 496.
84. Id. at para. 113.
“no”—because the Trial Chamber did not—but the reasons underlying those findings undermined the technical classification. For instance, a Trial Chamber might find an allegation about which the witness testified not proven but only because the witness’s testimony did not match the prosecution’s allegation. That is, the witness testified as to X, and even though the Trial Chamber may have believed the witness as to X, it did not matter because the existence of X did not support allegation Y that the prosecution was seeking to prove. These allegations generated a “no” entry because, technically speaking, the Trial Chamber did not rely on the witness’s testimony to prove the allegation, but I included explanatory notes to make clear that the Trial Chamber’s failure to rely on the witness’s testimony did not stem from concerns about the witness’s credibility or the reliability of her testimony but rather stemmed from a mismatch between the testimony and the allegation. The converse situation also regularly arose with respect to “yes” entries. Here, most commonly, a Trial Chamber would find proven the allegation about which the witness testified, but it would make clear that it did so only because the witness’s testimony was corroborated by other (ostensibly more credible or reliable) testimony. So, this allegation would generate a “yes” entry because, technically speaking, the Trial Chamber did rely on the witness’s testimony to prove an allegation, but my comments would provide valuable contextual information about the nature of the “yes.”

The explanatory information was important because I next considered, as a whole, all of the entries for all of the allegations for a particular witness. Taking all of the entries and explanations into account, I categorized each witness with one of the following four labels: “yes,” “no,” “yes and no,” and “?”.

85. See, e.g., Hategekimana Judgment, supra note 59, at para. 230 (finding that even though defendant provided witness with a laissez-passer, “the Chamber heard no evidence that the Accused also issued laissez-passer to soldiers, Interahamwe, armed civilians or any other members of an alleged joint criminal enterprise”); Muhimana Judgment, supra note 58, at paras. 60–61 (finding that the witness’s testimony fell “outside the scope” of the Indictment); Niyitegeka Judgment, supra note 15, at paras. 273, 287 (finding that even though witness saw the Accused, heard a gunshot, and saw two people dead, the witness did not see who fired the gun or where the gunshot came from, so “there is insufficient evidence in support of the allegation that the Accused killed the man and woman”); Setako Judgment, supra note 59, at paras. 409–410 (finding that there was widespread looting in Kigali but there was no “admissible evidence of Setako ordering or encouraging militiamen or soldiers in Kigali who committed looting”).

86. See, e.g., Gatere Judgment, supra note 58, at para. 341 (finding testimony of Witnesses BBJ and BCS “consistent and compelling . . . to the extent that they are corroborated”); Kanyarukiga Judgment, supra note 66, at para. 440 (noting that “Witnesses CBK, CBT, CDK and YAU are regarded with caution by the Chamber, particularly without corroboration” but their testimonies are corroborated by three other witnesses “regarded as credible”); Munyakazi Judgment, supra note 55, at para. 415 (finding testimony by Witnesses MM and MP identifying the Accused to be reliable though based on hearsay “because Munyakazi’s role is corroborated by Witnesses LCQ and BWW”).

87. See, e.g., Kanyarukiga Judgment, supra note 66, at para. 451 (finding, with respect to the Accused’s participation in a criminal enterprise, “there is no corroborated evidence of the Accused’s participation in a criminal enterprise”).

88. See, e.g., Kanyarukiga Judgment, supra note 66, at para. 452 (noting that “the Chamber finds the witness’s testimony that the Accused participated in a criminal enterprise . . . uncorroborated but . . . consistent with other allegations”).
tions; or (3) some of the allegations relevant to the witness were classified as “no,” but they were technical “no’s” as described above and did not reflect any Trial Chamber concerns about the witness’s credibility or the reliability of her testimony. Similarly, I classified witnesses as a “no” if one of the following three situations existed: (1) the Trial Chamber did not rely on the witness’s testimony and rejected all of the relevant allegations; (2) the Trial Chamber did not rely on the witness’s testimony and rejected the overwhelming majority of relevant allegations; or (3) some of the allegations relevant to the witness were classified as “yes,” but they were technical “yeses” as described above and did not reflect the Trial Chamber’s positive assessment of the witness’s credibility or the reliability of her testimony. I classified a witness as “yes and no” when the Trial Chambers relied on some but not all of the witness’s testimony to prove allegations. Finally, and very rarely, I classified a witness as “?” when the Trial Chamber simply did not provide enough information for me to determine whether they had relied on the witness’s testimony in proving or rejecting an allegation. I classified only two witnesses of the 342 in the dataset as “?”.

6. Other Data

My remaining data-gathering needs little explanation. I gathered data about the witnesses’ gender, ethnicity, imprisonment status, and accomplice status from the transcripts and judgments. I calculated the time between the subject testimony and the pre-trial statements/testimonies in months, and I rounded to the nearest month. Thus, I classified a statement dated between the first and fifteenth day of a month as occurring during that month, whereas I classified a statement occurring on the sixteenth day or later as occurring during the following month.

III. Serious Inconsistencies: The Who, What, Where, and When of Testimony that Diverges from Previous Representations

A. The Incidence of Serious Inconsistencies

Inconsistencies appearing in ICTR testimony pertain to a wide range of topics including facts about the crime itself, facts about the defendant’s participation in the crime, and facts relating to the witness’s observation of the events in question.87 Part II.A explained the significance of inconsistencies in international criminal fact-finding. Part III will provide a comprehensive picture of those inconsistencies, starting with their incidence. The data reveal that 67 percent of witnesses in the dataset testified in a way that was inconsistent to some degree with the witnesses’ previous statements/testimo-

nies. When we consider only inconsistencies that are serious, the percentage declines, though not dramatically so. The available transcripts and judgments show that 42 percent of witnesses testify in a way that is seriously inconsistent with their pre-trial statements/testimonies. When I adjusted the data to account for testimony held in camera,\(^8\) the percentage of witnesses who testified seriously inconsistently rose to 48 percent.

**B. Variation in Serious Inconsistencies Over Time**

Graph 1 depicts the incidence of serious inconsistencies over time. Both the graph and a logit regression showed no statistically significant time trend, so we must assume that the incidence of serious inconsistencies remained constant throughout the ICTR's life.\(^9\)

**GRAPH 1: SHOWING CHRONOLOGICAL INCIDENCE OF SERIOUS INCONSISTENCIES**

In order to isolate factors that might predict serious inconsistencies, I gathered a large quantity of data about the 342 witnesses in my dataset.\(^90\) I coded witnesses by gender, for instance, and learned that roughly equal proportions of men and women testify seriously inconsistently (38 percent for women and 44 percent for men). There is a slightly greater difference in the incidence of serious inconsistencies between the two ethnic groups (41 per-

---

88. For an explanation of that adjustment, see *supra* Part II.B.3.(b).
89. See Logit Regression 1: Explaining Serious Inconsistencies by Comprehensive Set of Explanatory Variables, *infra* app. 1.
90. As noted in Part II.B.3.b., all of the findings discussed henceforth in this Part are derived from the understated percentage of serious inconsistencies.
percent for Tutsi and 48 percent for Hutu), but it is not as substantial a difference as one might expect. Combining gender and ethnicity, however, did reveal an interesting disparity. In particular, whereas the incidence of serious inconsistencies among male Hutu witnesses was roughly equal to that of male Tutsi witnesses (47 percent for male Hutu compared to 44 percent for male Tutsi), a far greater proportion of female Hutu witnesses testified seriously inconsistently (71 percent) than female Tutsi witnesses (36 percent).

When we take additional factors in account, however, some of the differences I have just described disappear. A logit regression seeking to explain variation in serious inconsistencies that included not only gender and ethnicity but also a series of other potentially relevant factors,\(^9\) shows no statistically significant relationship between gender and serious inconsistencies or between ethnicity and serious inconsistencies. However, the regression does confirm that Hutu women are substantially more likely to testify seriously inconsistently than witnesses of any other gender-ethnic combination, and this result is statistically significant. The sample size of Hutu women witnesses is quite small,\(^9\) however, so this finding may be somewhat imprecise.

D. Serious Inconsistencies by Imprisonment and Accomplice Status

The tabular data suggests a correlation between a witness’s accomplice or imprisonment status and his likelihood of testifying seriously inconsistently. Specifically, whereas 60 percent of accomplices testified seriously inconsistently, only 40 percent of non-accomplices did so. However, the logit regression,\(^9\) which factored in a comprehensive set of explanatory variables, showed that accomplice status had no effect on the probability of a witness testifying seriously inconsistently. The tabular data shows even greater apparent divergences among witnesses based on their imprisonment status. I classified witnesses into three groups: (1) witnesses who had never been imprisoned for genocide crimes; (2) witnesses who had been imprisoned for genocide crimes but had been released; and (3) witnesses who had been imprisoned for genocide crimes and were still in prison at the time they testified. The data showed that only 38 percent of witnesses who were never imprisoned testified seriously inconsistently whereas 52 percent of witnesses imprisoned during trial did. Moreover, a whopping 75 percent of witnesses who were imprisoned but released before trial testified seriously inconsis-

---

91. These additional factors are the witnesses' accomplice status, imprisonment status, the number of pretrial statements/testimonies provided by each witness, and the starting date of the trial.
92. Of 342 witnesses, only seven were Hutu women.
93. Logit regressions differ from linear regressions because in a logit regression the dependent variable that a modeler is trying to explain is a percentage between zero and one. Regular linear regression ("ordinary least squares," or OLS) cannot guarantee that the predicted value of the dependent variable will fall between zero and one for all values of the explanatory variables. Logit regressions use a special, non-linear function form to achieve this end, and must be estimated with maximum likelihood methods instead of OLS.
tently. Again, however, the logit regression showed that when we include other relevant explanatory variables, a witness's imprisonment status had no statistically significant effect on the likelihood of finding a serious inconsistency in his testimony. 94

E. Serious Inconsistencies by Number and Type of Pretrial Statement/Testimony

Table 6 shows the relationship between the number of statements/testimonies witnesses provided and the percentage of serious inconsistencies in the population of witnesses who provided that number of statements/testimonies.

**TABLE 6: PROBABILITY OF SERIOUS INCONSISTENCIES BASED ON THE NUMBER OF PRE-TRIAL STATEMENTS/TESTIMONIES PROVIDED BY A WITNESS**

<table>
<thead>
<tr>
<th>Number of Pre-trial Statements/Testimonies Provided by a Witness</th>
<th>% of Witnesses with that Number of Statements/Testimonies who Testified Seriously Inconsistently</th>
<th>Number of Witnesses in the Dataset who Provided that Number of Statements/Testimonies</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24%</td>
<td>172</td>
</tr>
<tr>
<td>2</td>
<td>48%</td>
<td>77</td>
</tr>
<tr>
<td>3</td>
<td>62%</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>67%</td>
<td>18</td>
</tr>
<tr>
<td>5</td>
<td>75%</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>79%</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>100%</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>33%</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>100%</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>100%</td>
<td>1</td>
</tr>
</tbody>
</table>

As Table 6 indicates, an increased number of pretrial statements/testimonies is correlated with a dramatically increased incidence of serious inconsistencies. Indeed, merely increasing pretrial statements/testimonies from one to two doubles the proportion of witnesses who testify seriously inconsistently. Admittedly, the number of witnesses who provided more than six statements/testimonies is small, so that could reduce our confidence in that

94. See Logit Regression 1: Explaining Serious Inconsistencies by Comprehensive Set of Explanatory Variables, *infra* app. 1.
particular result, but the logit regression shows a highly statistically significant effect for the number of documents. In particular, the regression shows that for each of the first four statements or testimonies, adding a statement/testimony raises the probability of a serious inconsistency by approximately 11 percent. Adding more statements/testimonies continues to increase the probability of serious inconsistencies, but to a lesser degree as the number of statements/testimonies increases.\footnote{In particular, adding a fifth document for a witness who already has four raises the probability of a serious inconsistency by 9 percent. Adding a sixth raises the probability by 7 percent; adding a seventh raises the probability by 5 percent, and adding an eighth raises the probability by 3 percent. According to the tabular data, witnesses who have provided nine or more statements or testimonies are 100 percent likely to testify seriously inconsistently with one of them. See Marginal Effects Calculation 1: Effects of the Number of Statements/Testimonies on Serious Inconsistencies as Estimated by Logit Regression 1, infra app. 1.}

Overall, 33 percent of statements/testimonies gave rise to serious inconsistencies, but these were not evenly distributed across the different kinds of statements/testimonies. For instance, statements/testimonies submitted to foreign courts generated the smallest proportion of serious inconsistencies whereas statements/testimonies submitted to Rwandan courts generated the largest. The specific percentages are shown in Table 7 below.

**Table 7: Serious Inconsistencies by Document Type**

<table>
<thead>
<tr>
<th>Type of Document</th>
<th>Percentage of those Documents Found Seriously Inconsistent with ICTR Testimony in the Subject Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rwandan Court Statements/Testimonies</td>
<td>42%</td>
</tr>
<tr>
<td>Gacaca Statements/Testimonies</td>
<td>35%</td>
</tr>
<tr>
<td>ICTR Pre-Trial Statements</td>
<td>31%</td>
</tr>
<tr>
<td>Testimony in Previous ICTR Cases</td>
<td>30%</td>
</tr>
<tr>
<td>Foreign court Statements/Testimonies</td>
<td>22%</td>
</tr>
<tr>
<td>All Documents Combined</td>
<td>33%</td>
</tr>
</tbody>
</table>

I next considered whether the incidence of serious inconsistencies in the different types of statements/testimonies had changed during the lifespan of the ICTR. I did not consider potential time trends for foreign court statements/testimonies—because the sample size was too small—or for gacaca statements/testimonies—because gacaca proceedings were not fully underway until many of the trials in my dataset had already concluded. However, Graph 2 shows the chronological incidence of serious inconsistencies with ICTR pre-trial statements, ICTR previous testimony, and Rwandan court statements/testimonies. A simple regression of serious inconsistency by document type over time suggested no statistically significant time trend for

\footnote{The dataset included only nine pre-trial statements/testimonies from foreign courts, so this finding is less precise than the others.}
testimony that was seriously inconsistent with ICTR pre-trial statements. However, the regression did suggest a statistically significant downward trend for testimony that was seriously inconsistent with previous ICTR testimony and a statistically significant upward trend for testimony that was seriously inconsistent with Rwandan court statements/testimonies. Specifically, the regressions suggested that the percentage of serious inconsistencies in previous ICTR testimony decreased by an average of 2.6 percent per year while the percentage of serious inconsistencies in Rwandan court statements/testimonies increased by an average of 4 percent per year. As already noted, the data shows no statistically significant time trend for serious inconsistencies across the whole dataset, but it does reveal subcategory trends that offset one another.

**GRAPH 2: SHOWING THE PERCENTAGE OF SERIOUS INCONSISTENCIES IN DIFFERENT TYPES OF STATEMENTS/TESTIMONIES OVER TIME**

Finally, I considered whether the length of time between the pretrial statement/testimony and the subject testimony predicted serious inconsistencies. I expected that witnesses were more likely to testify seriously incon-

---

97. See Simple Regression 1: Serious Inconsistencies Over Time, by Document Types, infra app. 1.
98. Id.
99. Note that there is missing data for some document types at some trials. For example, in Graph 2, we can see that for many trials there were no witnesses who had previously testified in an ICTR case. Such missing data is the reason that a number of the graphs in the remainder of this Article have missing stretches and isolated markers.
sistently with old statements/testimonies than with more recent ones. The data does not bear out that expectation, however. Table 8 and Graph 3 show the probability of serious inconsistencies for statements/testimonies that were provided over various time periods. The data shows that ICTR testimony is less likely to seriously diverge from very recent statements/testimony, but after that, the probability of serious inconsistencies does not increase as more time elapses between statements/testimonies and the subject testimony.

**Table 8: Relationship between Serious Inconsistencies and the Length of Time Between the Statement/Testimony and the Subject Testimony**

<table>
<thead>
<tr>
<th>Length of Time Between Subject Testimony and Pre-Trial Statement/Testimony</th>
<th>Number of Statements/Testimonies in this Time Period with which Subject Testimony Is Seriously Inconsistent</th>
<th>Total Number of Statements/Testimonies in this Time Period</th>
<th>% of Statements/Testimonies in this Time Period with which Subject Testimony Is Seriously Inconsistent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than one year</td>
<td>11</td>
<td>47</td>
<td>23%</td>
</tr>
<tr>
<td>1–2 years</td>
<td>20</td>
<td>54</td>
<td>37%</td>
</tr>
<tr>
<td>2–3 years</td>
<td>25</td>
<td>48</td>
<td>52%</td>
</tr>
<tr>
<td>3–4 years</td>
<td>17</td>
<td>39</td>
<td>44%</td>
</tr>
<tr>
<td>4–5 years</td>
<td>22</td>
<td>60</td>
<td>37%</td>
</tr>
<tr>
<td>5–6 years</td>
<td>23</td>
<td>45</td>
<td>51%</td>
</tr>
<tr>
<td>6–7 years</td>
<td>21</td>
<td>46</td>
<td>46%</td>
</tr>
<tr>
<td>7–8 years</td>
<td>16</td>
<td>38</td>
<td>42%</td>
</tr>
<tr>
<td>8–9 years</td>
<td>17</td>
<td>39</td>
<td>44%</td>
</tr>
<tr>
<td>9–10 years</td>
<td>38</td>
<td>97</td>
<td>39%</td>
</tr>
<tr>
<td>10–12 years</td>
<td>20</td>
<td>44</td>
<td>45%</td>
</tr>
<tr>
<td>13+ years</td>
<td>11</td>
<td>47</td>
<td>23%</td>
</tr>
</tbody>
</table>
IV. ASSESSING TESTIMONY: A COMPREHENSIVE EXAMINATION OF THE ICTR'S CREDIBILITY AND RELIABILITY DETERMINATIONS

Reading numerous ICTR judgments has led me to expect that certain factors would be statistically significant predictors of Trial Chambers' credibility and reliability findings. As this section reveals, the data supports some of my expectations but by no means all. Table 9 begins the discussion by displaying the percentages of witnesses across the entire dataset whom the Trial Chambers found wholly or partially credible and on whose testimony the Trial Chambers wholly or partially relied.

TABLE 9: TRIAL CHAMBERS' CREDIBILITY AND RELIABILITY ASSESSMENTS

<table>
<thead>
<tr>
<th>Credibility</th>
<th>Relied Upon</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Witnesses Found Totally Credible</td>
<td>% of Witnesses % of Witnesses Whose Testimony was Totally Relied Upon</td>
</tr>
<tr>
<td>% of Witnesses Found Totally or Partially Credible</td>
<td>% of Witnesses Whose Testimony was Totally Relied Upon</td>
</tr>
<tr>
<td>55%</td>
<td>61%</td>
</tr>
<tr>
<td>37%</td>
<td>57%</td>
</tr>
</tbody>
</table>

A. The Trial Chambers' Assessments of Witness Testimony over Time

This section reports one of this Article's most notable findings: that the Trial Chambers' willingness to find prosecution witnesses credible and to rely on their testimony declined significantly over time. This finding comes
as no surprise to anyone who has read many ICTR judgments. Early judgments feature fewer discussions of witness credibility and reliability than later judgments, and what discussions there are in the early judgments are less detailed and in-depth. In addition, in early cases, Trial Chambers seemed more likely to explain away serious inconsistencies and other testimonial deficiencies than they did in later cases. The data confirms my impressionistic sense that Trial Chambers treated prosecution witness testimony more skeptically as time passed.

Graph 4 depicts the Trial Chambers' positive total and partial credibility findings over time, and Graph 5 depicts the Trial Chambers' willingness to totally and partially rely on prosecution testimony over time.

**GRAPH 4: CHRONOLOGICALLY SHOWING PERCENTAGE OF WITNESSES TRIAL CHAMBERS FOUND WHOLLY OR PARTIALLY CREDIBLE**

- Percentage of witnesses found to be wholly credible
- Percentage of witnesses found to be wholly or partially credible
A simple regression of credibility on time confirms a statistically significant downward time trend. In particular, simple regressions suggest that the Trial Chambers' positive total credibility findings declined on average by 2.1 percent per year, and their positive total or partial credibility findings declined on average by 2.0 percent per year. Similarly, the simple regressions suggest that the Trial Chambers' willingness to totally rely on a prosecution witness's testimony declined by an average of 2.4 percent per year, whereas their willingness to totally or partially rely on a prosecution witness's testimony declined by an average of 2.1 percent per year. In an effort to account for the effect of other factors, I included a timing variable in a logit regression featuring a comprehensive set of other explanatory variables. It confirms the statistically significant negative time trend, showing that if we fix other variables at their average values, the passage of time, isolated, reduced by 51 percent the likelihood that a Trial Chamber would

---

100. See Simple Regression 2: Trial Chambers' Findings of Total Credibility Over Time, infra app. 2. For all of the simple regressions, the calculations are based on the mean value of independent variables.
101. See Simple Regression 3: Trial Chambers' Findings of Total and Partial Credibility Over Time, infra app. 2.
102. See Simple Regression 4: Trial Chambers' Total Reliance on Testimony Over Time, infra app. 2.
103. See Simple Regression 5: Trial Chambers' Total and Partial Reliance on Testimony Over Time, infra app. 2.
find a witness totally credible,\textsuperscript{104} and it reduced by 46 percent the likelihood that a Trial Chamber would find a witness totally or partially credible.\textsuperscript{105} Put another way, during the first trial in the dataset, a witness, who was average in every other respect, had a 97 percent chance of being found totally credible and a 99 percent chance of being found totally or partially credible. By the last trial in the dataset, that same witness had only a 46 percent chance of being found totally credible and a 53 percent chance of being found totally or partially credible. The logit regression produced similar findings regarding the Trial Chambers’ willingness to rely on prosecution witness testimony. That is, if we fix all other variables at their average values, the passage of time, isolated, reduced the Trial Chambers’ willingness to wholly rely on a witness’s testimony by 58 percent\textsuperscript{106} and to wholly or partially rely on the witness’s testimony by 51 percent.\textsuperscript{107} That is, during the first trial in the dataset, a witness, who was average in every other respect, had an 81 percent chance of having her testimony wholly relied upon, and a 97 percent chance of having her testimony totally or partially relied upon. By the last trial in the dataset, that same witness had only a 23 percent chance of having her testimony totally relied upon and a 46 percent chance of having her testimony totally or partially relied upon.

\textbf{B. The Trial Chambers’ Assessment of Witness Testimony by Gender and Ethnicity}

The tabular data shows that Trial Chambers found a higher proportion of female witnesses credible than male, and that they were slightly more willing to rely on the testimony of female witnesses. Table 10 shows the comparable proportions. The divergence between the genders is greatest when we consider those witnesses whom the Trial Chamber found wholly credible or whose testimony the Trial Chamber wholly relied upon. Indeed, when we include witnesses who are found partially credible or whose testimony is partially relied upon, the divergence in the Trial Chambers’ treatment of the two genders narrows substantially, in the case of credibility, and disappears entirely, in the case of reliance.

\textsuperscript{104} See Logit Regression 2: Trial Chambers’ Total Credibility Findings by Comprehensive Set of Explanatory Variables, infra app. 3.

\textsuperscript{105} See Logit Regression 3: Trial Chambers’ Total and Partial Credibility Findings by Comprehensive Set of Explanatory Variables, infra app. 3.

\textsuperscript{106} See Logit Regression 4: Trial Chambers’ Total Reliance on Testimony by Comprehensive Set of Explanatory Variables, infra apps. 3–4.

\textsuperscript{107} See Logit Regression 5: Trial Chambers’ Total and Partial Reliance on Testimony by Comprehensive Set of Explanatory Variables, infra app. 5.
Turning next to ethnicity, tabular data suggests that Trial Chambers regarded Tutsi witnesses as more credible and that Trial Chambers were more likely to rely on their testimony, as Table 11 shows. As with gender, the ostensible disparities decreased when findings of partial credibility and partial reliance were included, but some disparity remained.

### TABLE 11: TRIAL CHAMBERS' CREDIBILITY AND RELIANCE FINDINGS BY ETHNICITY

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Credibility</th>
<th>Relied Upon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Witnesses Found Wholly Credible</td>
<td>% of Witnesses Found Wholly or Partially Credible</td>
</tr>
<tr>
<td>Hutu</td>
<td>45%</td>
<td>53%</td>
</tr>
<tr>
<td>Tutsi</td>
<td>60%</td>
<td>65%</td>
</tr>
</tbody>
</table>

Combining gender and ethnicity compounded whatever divergence appeared when we compared the characteristics separately. That is, the tabular data above indicates that Trial Chambers were slightly more likely to find female witnesses credible and to rely on their testimony than male witnesses; it likewise indicated that Trial Chambers were slightly more likely to find Tutsi witnesses credible and to rely on their testimony than Hutu witnesses. Therefore, Table 12 unsurprisingly shows that Trial Chambers are most likely to find credible and to rely on the testimony of Tutsi female witnesses and least likely to find credible and to rely on the testimony of Hutu male witnesses. That said, the differences are relatively small. Moreover, when I factored in other relevant explanatory variables in a logit regression, the
effect disappeared. The regression revealed no statistically significant effect for gender, ethnicity, or the gender-ethnic combination.108

Table 12: Trial Chambers’ Credibility and Reliance Findings by Gender-Ethnic Combinations

<table>
<thead>
<tr>
<th>Gender-Ethnic Combinations</th>
<th>Credibility</th>
<th>Relied Upon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Witnesses Found Wholly Credible</td>
<td>% of Witnesses Whose Testimony was Wholly Relied Upon</td>
</tr>
<tr>
<td>Hutu Female</td>
<td>57%</td>
<td>29%</td>
</tr>
<tr>
<td>Tutsi Female</td>
<td>64%</td>
<td>45%</td>
</tr>
<tr>
<td>Hutu Male</td>
<td>44%</td>
<td>29%</td>
</tr>
<tr>
<td>Tutsi Male</td>
<td>58%</td>
<td>41%</td>
</tr>
</tbody>
</table>

Finally, I considered what role, if any, gender or ethnicity played in the Trial Chambers’ credibility or reliability determinations over time. We know that the Trial Chambers’ positive credibility and reliability findings declined over time for the whole population of witnesses in the dataset, but here I sought to determine if that decline varied by gender or ethnicity. Graphs 6 through 13 provide a chronological depiction of the Trial Chambers’ credibility and reliability findings by gender and ethnicity. Graphs 6 through 9 show the Trial Chambers’ total and partial credibility findings by gender followed by the Trial Chambers’ total and partial reliability findings by gender. Graphs 10 through 13 show the same variables by ethnicity.

108. See Logit Regressions 2, 3, 4, and 5, infra apps. 3–5.
Graph 6: Showing Trial Chambers' Total Credibility Findings by Gender Over Time

Graph 7: Showing Trial Chambers' Total and Partial Credibility Findings by Gender Over Time
Graph 8: Showing Trial Chambers' Total Reliance on Witness Testimony by Gender Over Time

Graph 9: Showing Trial Chambers' Total or Partial Reliance on Witness Testimony by Gender Over Time
Graph 10: Showing Trial Chambers' Total Credibility Findings by Ethnicity Over Time

Graph 11: Showing Trial Chambers' Total or Partial Credibility Findings by Ethnicity Over Time
Graph 12: Showing Trial Chambers' Total Reliance on Witness Testimony by Ethnicity Over Time

Graph 13: Showing Trial Chambers' Total and Partial Reliance on Witness Testimony by Ethnicity Over Time
The graphs show what appears to be unexplained variation in the sample, but some simple regressions suggest several interesting time trends. In particular, the regressions examined the effect of gender on the four different categories of Trial Chambers' findings over time (that also appear in the graphs):

1. Witnesses found totally credible;
2. Witnesses found totally or partially credible;
3. Witnesses whose testimony was totally relied upon; and
4. Witnesses whose testimony was totally or partially relied upon.

The most interesting results relate to witnesses in category 1, those found totally credible. There, I found a substantial—and statistically significant—difference in the Trial Chambers' findings over time by gender. In particular, the simple regression suggested that, although the Trial Chambers' willingness to find male witnesses totally credible declined by an average of 2.8 percent per year, it stayed virtually the same for female witnesses.\textsuperscript{109} Simple regressions on the other three categories also suggested a similar—though smaller—divergence between male and female witnesses; this conclusion is less certain, as the findings for the female witnesses are not statistically significant. With respect to male witnesses, the regression showed a statistically significant decline in the Trial Chambers' willingness to find them totally or partially credible and to rely on their testimony in whole or in part. The average decline is fairly similar across each of the three categories: between 2.3 and 2.5 percent per year.\textsuperscript{110} With respect to female witnesses, there appeared to be no decline in the Trial Chambers' willingness to find witnesses totally or partially credible, whereas the regression suggested that the Trial Chambers' categories 3 and 4 findings for female witnesses declined by 1.7 percent and 1.3 percent per year, respectively.\textsuperscript{111} As noted, however, the findings of categories 2, 3, and 4 for female witnesses were not statistically significant.

The regressions also showed divergences in the Trial Chambers' treatment of the testimony of the two ethnic groups over time. In particular, the Trial Chambers' willingness to credit Hutu witnesses and to rely on their testimony declined much more substantially than the concomitant decline for Tutsi witnesses. The regressions showed the greatest divergence in the Trial Chambers' credibility findings. In particular, the Trial Chambers' positive total credibility findings and their positive total and partial credibility findings declined by an average of 4.4 percent and 4.5 percent per year, respectively, for Hutu witnesses whereas they declined by an average of only 1.1

\textsuperscript{109} Summary 1: Simple Regression Statistics for Credibility Findings Over Time by Gender, infra app. 6.
\textsuperscript{110} Id.; Summary 2: Simple Regression Statistics for Reliance Findings Over Time by Gender, infra app. 6.
\textsuperscript{111} Summary 1, supra note 109; Summary 2, supra note 110.
percent and 1 percent per year, respectively, for Tutsi witnesses. The Trial Chambers' reliance on witness testimony also declined over time at a greater rate for Hutu witnesses, but these divergences were not as substantial or as certain.

C. The Trial Chambers' Assessment of Witness Testimony by Accomplice Status and Imprisonment Status

One gets the impression from ICTR judgments that Trial Chambers cast a skeptical eye on the testimony of witnesses who were accomplices and witnesses who have been imprisoned for genocide crimes. Trial Chambers frequently note the various motivations that accomplice or imprisoned witnesses have to falsely inculpate defendants, and they frequently claim to treat such witnesses' testimony "with caution." These sorts of comments could lead a reader to believe that Trial Chambers were less likely to find accomplices and imprisoned witnesses credible and less likely to rely on their testimony. The tabular data is in keeping with that assumption. As Table 13 shows, when it comes to witnesses who were never imprisoned for genocide crimes, the Trial Chambers found them credible and relied on their testimony in much higher proportions than they did witnesses who have been imprisoned for genocide crimes. Indeed, Trial Chambers totally relied on the testimony of Hutu witnesses and its total or partial reliance on their testimony declined by a statistically significant average of 2.7 percent and 3.4 percent per year, respectively. The corresponding declines for Tutsi witnesses appeared to be 1.4 percent and 1.1 percent per year, though these were not statistically significant. Id.

114. See, e.g., Nizeyimana Judgment, supra note 59, at paras. 111, 413, 441, 504, 559-560, 607, 820-821, 836-837, 1110, 1138-1139; Nzabonimana Judgment, supra note 58, at paras. 1276, 1344, 1348, 1480; Kanyarukiga Judgment, supra note 66, at paras. 468, 576, 578.


113. Summary 4: Simple Regression Statistics for Reliance Findings Over Time by Ethnicity, infra app. 6. Specifically, the Trial Chambers' total reliance on the testimony of Hutu witnesses and its total or partial reliance on their testimony declined by a statistically significant average of 2.7 percent and 3.4 percent per year, respectively. The corresponding declines for Tutsi witnesses appeared to be 1.4 percent and 1.1 percent per year, though these were not statistically significant. Id.
witnesses imprisoned for genocide crimes. In particular, Trial Chambers are approximately twice as likely to find non-accomplice witnesses credible and to rely on their testimony as they are accomplice witnesses. Again, however, the disparity narrowed when witnesses found to be partially credible or whose testimony is partially relied upon were included.

**Table 13: Trial Chambers' Credibility and Reliance Findings by Imprisonment Status and Accomplice Status**

<table>
<thead>
<tr>
<th>Imprisonment Status and Accomplice Status</th>
<th>Credibility</th>
<th>Relied Upon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Witnesses Found Wholly Credible</td>
<td>% of Witnesses Found Wholly or Partially Credible</td>
</tr>
<tr>
<td>Never Imprisoned for Genocide</td>
<td>60%</td>
<td>65%</td>
</tr>
<tr>
<td>Imprisoned for Genocide but Released Before Trial</td>
<td>35%</td>
<td>40%</td>
</tr>
<tr>
<td>Imprisoned for Genocide and Still Detained at Trial</td>
<td>39%</td>
<td>50%</td>
</tr>
<tr>
<td>Not an Accomplice to the Defendant</td>
<td>61%</td>
<td>65%</td>
</tr>
<tr>
<td>Accomplice to the Defendant</td>
<td>34%</td>
<td>46%</td>
</tr>
</tbody>
</table>

Given this tabular data, along with the Trial Chambers' skeptical comments about the testimony of accomplice and imprisoned witnesses, I expected that a regression would confirm that these two characteristics had a statistically significant effect on the Trial Chambers' findings. However, the logit regression took account of a comprehensive set of variables and did not show a statistically significant effect either for accomplice status or imprisonment status.116

---

116. See Logit Regressions 2, 3, 4, and 5, infra apps. 3-5.
I next sought to ascertain whether imprisonment status may have influenced the Trial Chambers' credibility and reliability decisions over time. Graphs 14 through 17 chronologically depict the tabular data on that question.

**Graph 14: Chronologically Showing Trial Chambers' Total Credibility Findings by Imprisonment Status**

- Never Imprisoned
- Imprisoned but Released
- Imprisoned at Trial
Graph 15: Showing Trial Chambers' Total and Partial Credibility Findings by Imprisonment Status Over Time

Graph 16: Showing Trial Chambers' Total Reliance on Witness Testimony by Imprisonment Status Over Time
Simple regressions reveal some interesting time trends. I categorized the Trial Chambers' findings into the same four categories discussed above in subsection B, namely:

1. Witnesses found totally credible;
2. Witnesses found totally or partially credible;
3. Witnesses whose testimony was totally relied upon; and
4. Witnesses whose testimony was totally or partially relied upon.

I considered those findings with respect to the following three classes of witnesses:

A. Witnesses never imprisoned for genocide;
B. Witnesses imprisoned for genocide but released before trial; and
C. Witnesses imprisoned for genocide and still detained at trial.

Table 14 presents the results of the simple regressions. It shows an overall decline in the Trial Chambers' positive credibility and reliability findings over time for all three categories of witnesses. However, declines for the two groups of imprisoned witnesses were much more substantial than for the never-imprisoned witnesses. We cannot be as certain about the findings regarding the never-imprisoned witnesses because some of them were not statistically significant, but the data does clearly show (1) that the Trial Chambers' confidence in the testimony of imprisoned witnesses declined dramatically during the course of the ICTR's life; and (2) that its confidence in the testimony of never-imprisoned witnesses did not change nearly as
Finally, the most notable finding, perhaps, stems from a comparison of the Trial Chambers' treatment of those who were imprisoned but released and those who remained imprisoned at the time of their testimony. That comparison shows that the decline in the Trial Chambers' positive credibility and reliance findings was—counterintuitively—much steeper for the witnesses who were imprisoned but released before trial than for witnesses who were still detained at trial. Part V will discuss this surprising result in more detail.

---

117. Although some of the findings with respect to never-imprisoned witnesses were not statistically significant, they were close. More importantly, the 95 percent confidence intervals surrounding these findings provide assurance that the Trial Chambers' treatment of the testimony of never-imprisoned witnesses did not change at the rate that it did for imprisoned witnesses.
<table>
<thead>
<tr>
<th>Imprisonment Status</th>
<th>Average Decline per Year in Trial Chambers' Willingness to Find Witnesses Totally Credible</th>
<th>Statistical Significance of the Totally Credible Finding</th>
<th>Average Decline per Year in Trial Chambers' Willingness to Find Witnesses Totally or Partially Credible</th>
<th>Statistical Significance of the Totally or Partially Credible Finding</th>
<th>Average Decline per Year in Trial Chambers' Willingness to Totally or Partially Rely on Witnesses' Testimony</th>
<th>Statistical Significance of the Total Reliance Finding</th>
<th>Statistical Significance of the Total or Partial Reliance Finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never Imprisoned for Genocide</td>
<td>-1%</td>
<td>Not Statistically Significant ($t$: -1.45)</td>
<td>-1%</td>
<td>Not Statistically Significant ($t$: -1.23)</td>
<td>-1.6%</td>
<td>Statistically Significant ($t$: -2.23)</td>
<td>Not Statistically Significant ($t$: -1.43)</td>
</tr>
<tr>
<td>Imprisoned for Genocide but Released before Trial</td>
<td>-10.7%</td>
<td>Statistically Significant ($t$: -2.12)</td>
<td>-10.6%</td>
<td>Statistically Significant ($t$: -2.14)</td>
<td>-5%</td>
<td>Not Statistically Significant ($t$: -0.7)</td>
<td>Statistically Significant ($t$: -2.1)</td>
</tr>
<tr>
<td>Imprisoned for Genocide and Still Detained at Trial</td>
<td>-7.7%</td>
<td>Statistically Significant ($t$: -2.91)</td>
<td>-8%</td>
<td>Statistically Significant ($t$: -3.12)</td>
<td>-6.9%</td>
<td>Statistically Significant ($t$: -2.18)</td>
<td>Statistically Significant ($t$: -2.8)</td>
</tr>
</tbody>
</table>
D. The Interaction Between Credibility, Reliability, and Serious Inconsistencies

Because there is reason to believe that the testimonial deficiency most likely to impair accurate fact-finding is serious inconsistency, it is particularly important to probe the way in which the Trial Chambers' actual fact-finding is influenced by serious inconsistency. Subsection 1 considers the same question that we have been exploring in all of the sections in this Part, namely, the Trial Chambers' positive credibility and reliability findings. However, instead of categorizing our population of witnesses by gender or ethnicity, for example, subsection 1 categorizes the witnesses by whether their testimony contains a serious inconsistency or not. Next, subsection 2 asks a more targeted question: when are Trial Chambers willing to rely specifically on testimony that was seriously inconsistent with previous statements/testimonies in order to find a fact beyond a reasonable doubt? Finally, subsection 3 explores the public representation of the Trial Chambers' treatment of serious inconsistencies by examining their willingness to mention such inconsistencies in their judgments.

1. The Trial Chambers' Assessments of Witness Testimony with Serious Inconsistencies

Table 15 shows that a much higher proportion of witnesses without serious inconsistencies are found credible and reliable than witnesses with serious inconsistencies.

<table>
<thead>
<tr>
<th>Presence of Serious Inconsistency</th>
<th>Credibility</th>
<th>Relied Upon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of Witnesses Found Wholly Credible</td>
<td>% of Witnesses Found Wholly or Partially Credible</td>
</tr>
<tr>
<td>Witnesses with Serious Inconsistencies</td>
<td>41%</td>
<td>50%</td>
</tr>
<tr>
<td>Witnesses without Serious Inconsistencies</td>
<td>84%</td>
<td>89%</td>
</tr>
</tbody>
</table>
The results of the tabular data are confirmed by a logit regression that includes serious inconsistencies as one of a comprehensive set of explanatory variables. In particular, the regression revealed that, if all other variables are fixed at their average levels, witnesses who testify seriously inconsistently are 44 percent less likely to be found totally credible and 35 percent less likely to be found totally or partially credible than witnesses who do not testify seriously inconsistently. Similarly, Trial Chambers were 33 percent less likely to totally rely on the testimony of a witness with a serious inconsistency and 35 percent less likely to totally or partially rely on the testimony of a witness with a serious inconsistency.

Graphs 18 through 21 depict the Trial Chambers' credibility and reliability assessments over time for witnesses with and without serious inconsistencies.

**Graph 18: Showing Trial Chambers' Total Credibility Findings by Serious Inconsistencies Over Time**

![Graph](image-url)

---

118. See Marginal Effects of Serious Inconsistencies on the Trial Chambers' Total Credibility Findings, infra app. 3.
119. See Marginal Effects of Serious Inconsistencies on the Trial Chambers' Total and Partial Credibility Findings, infra app. 4.
120. See Marginal Effects of Serious Inconsistencies on the Trial Chambers' Total Reliance Findings, infra app. 5.
121. See Marginal Effects of Serious Inconsistencies on the Trial Chambers' Total and Partial Reliance Findings, infra app. 5.
Graph 19: Showing Chambers' Total and Partial Credibility Findings by Serious Inconsistency Over Time

Graph 20: Showing Trial Chambers' Total Reliance on Witness Testimony by Serious Inconsistency Over Time
Running some simple regressions allows us to compare the Trial Chambers' treatment of these two groups of witnesses over time, and they show dramatic differences when it comes to credibility findings. In particular, the regression suggests that the Trial Chambers' positive credibility findings for witnesses without serious inconsistencies stayed virtually the same through the course of the cases in the dataset, though that result is not statistically significant. At the same time, the regressions show a marked decline over time in the Trial Chambers' positive credibility findings for witnesses with serious inconsistencies. Specifically, the Trial Chambers' positive credibility findings, both total and partial, for witnesses with serious inconsistencies declined by an average of just under 5 percent per year.\textsuperscript{122}

The regressions examining the Trial Chambers' reliability findings for witnesses with and without serious inconsistencies present a murkier picture. For instance, the Trial Chambers' willingness to totally rely on witness testimony appeared to decline at an equal average rate of 1.9 percent per year for both sets of witnesses; however, that result is statistically significant only for witnesses without serious inconsistencies. When it comes to the Trial Chambers' willingness to rely in whole or in part on witness testimony, the same divergence seen in the Trial Chambers' credibility findings is again present. That is, the Trial Chambers' willingness to rely in whole or

\textsuperscript{122} Simple Regression Statistics for Credibility and Reliance Findings Over Time by Serious Inconsistencies in Testimony, infra app. 6.
part on witnesses with serious inconsistencies declined at a much greater rate, namely an average of 3.4 percent per year compared with an average of only 0.6 percent per year for witnesses without serious inconsistencies. However, only the result for witnesses with serious inconsistencies was statistically significant.123

2. Trial Chambers' Specific Reliance on Seriously Inconsistent Testimony

Throughout this Article, I have presented data on the Trial Chambers' reliance on witness testimony, but until now, when assessing a Trial Chamber's willingness to rely on a particular witness's testimony, I took account of the Trial Chamber's treatment of all of the witness's testimony. This section also examines the Trial Chambers' reliance on witness testimony, but it asks a more targeted question. For this section, I considered only witnesses whose testimony contained serious inconsistencies, and I asked: for what percentage of such witnesses did the Trial Chambers rely specifically on the seriously inconsistent testimony in order to find the relevant fact? Tabular data shows the answer to that question to be 33 percent.

Graph 22 below shows how the Trial Chambers' reliance on seriously inconsistent testimony has changed over time. The graph shows a clear decline in the Trial Chambers' willingness to rely on seriously inconsistent testimony, and a simple regression confirms that decline. In particular, the Trial Chambers' reliance on seriously inconsistent testimony declined by a statistically significant average of 1.4 percent per year.124

123. Id.
124. Simple Regression 6: Trial Chamber's Reliance on Testimony with Serious Inconsistency to Find Facts, infra app. 2.
Graph 22: Chronologically Showing Trial Chambers’ Reliance on Seriously Inconsistent Testimony

![Graph](image)

Percentage of Witnesses Whose Seriously Inconsistent Testimony Was Relied Upon to Prove Allegation

3. Trial Chambers’ References to Serious Inconsistencies in their Judgments

Although ICTR judgments are very lengthy and typically include a detailed description of each witness’s testimony, they do not always mention serious inconsistencies. In particular, my dataset reveals that Trial Chambers mention about 75 percent of serious inconsistencies in their judgments. Again, however, the data shows a fascinating time trend. As Graph 23 clearly shows, Trial Chambers’ references to serious inconsistencies increased over time. Indeed, a simple regression produced the statistically significant finding that references to serious inconsistencies increased an average of 1.6 percent per year.

---

125. ICTR judgments in single-defendant cases typically run at least 100 pages and often run several hundred pages.
126. Fact-Finding Without Facts, supra note 7, at 180 & n.796 (containing examples).
127. Specifically, Trial Chambers mentioned 72.22 percent of serious inconsistencies and failed to mention 21.53 percent. For the remaining 5.56 percent of witnesses with serious inconsistencies, the Trial Chamber mentioned at least one of the inconsistencies but failed to mention at least one.
128. Simple Regression 7: Trial Chamber’s References to Serious Inconsistencies in Judgments Over Time, infra app. 3.
The findings detailed in Parts III and IV provide interesting insights into international criminal fact-finding that I will explore here. Section A summarizes those findings and section B explains them.

A. Summarizing the Findings

I begin with the findings on serious inconsistencies. Although, as a general matter, I did very little quantifying in Fact-Finding Without Facts, I did make some effort to quantify serious inconsistencies. In particular, I examined six ICTR cases and three SCSL cases, and I found that, on average, approximately 50 percent of witnesses in those cases testified seriously inconsistently with their previous statements/testimonies. For this study, I constructed a far larger database. Yet, interestingly, it produced very similar results: 48 percent of prosecution witnesses testified seriously inconsistently. The incidence of serious inconsistencies varied, sometimes dramatically, with the case. Defense counsel in Muyakazi, for instance, failed even to mention witness statements for more than 70 percent of the prosecution witnesses; therefore, it came as no surprise that I identified serious inconsistencies in the testimony of only 9 percent of prosecution witnesses. By contrast, in Setako, Nzabonimana, and Ndahimana nearly three-quarters of prosecution witnesses' testimony contained serious inconsistencies. Although the case-by-case incidence of serious inconsistencies did vary, that

129. The three SCSL cases were the only SCSL cases in trial at the time I did my research, so my findings with respect to those cases encompassed all of the SCSL data then in existence.
variation did not produce a statistically significant increase or decrease in
the incidence of serious inconsistencies over the course of the ICTR's life.
However, the data does show time trends for some subcategories. For in-
stance, as time passed, the incidence of serious inconsistencies with previous
ICTR testimony decreased while the incidence of serious inconsistencies
with Rwandan court statements/testimonies increased. The incidence of seri-
ous inconsistencies with ICTR pre-trial statements remained constant.

The data also shows that a host of witness characteristics had no statisti-
cally significant effect on the likelihood that the witness would testify seri-
ously inconsistently. For instance, although tabular data showed that a
higher proportion of male witness testimony contained serious inconsisten-
cies than did female witness testimony and that a higher proportion of Hutu
witness testimony contained serious inconsistencies than did Tutsi witness
testimony, those effects disappeared in a regression that took account of
other relevant factors. When it came to the intersection of gender, ethnicity,
and serious inconsistencies, the only statistically significant finding was that
Hutu women were more likely to testify seriously inconsistently than any
other gender-ethnic combination. The fact that a witness had been an ac-
complice of the defendant or had been imprisoned for a genocide crime also
had no statistically significant effect on the likelihood of serious inconsisten-
cies in that witness's testimony. By contrast, some details surrounding the
previous statements/testimonies themselves did prove statistically signifi-
cant. In particular, the number of statements/testimonies that a witness pro-
vided was a very strong predictor of serious inconsistencies, though the
length of time between the statement/testimony and the subject testimony
was not. Finally, the data showed that witnesses were more likely to testify
seriously inconsistently with statements/testimonies that had been submit-
ted to Rwandan courts, and to a lesser extent gacaca courts, than with other
statements/testimonies.

When it comes to the Trial Chambers' assessments of witness testimony,
my research reveals two particularly notable, and heartening, findings. The
first is that Trial Chambers became less likely to find prosecution witnesses
credible or to rely on their testimony over the course of the ICTR's life. The
second is that Trial Chambers were considerably less likely to find witnesses
credible or to rely on their testimony when the testimony of those witnesses
contained serious inconsistencies. This latter finding is notable because no
other witness characteristic proved a statistically significant predictor of
Trial Chamber credibility and reliability determinations. Specifically, the
data showed no statistically significant effect for gender, ethnicity, gender
and ethnicity in combination, accomplice status, or imprisonment status. By
contrast, a witness whose testimony had a serious inconsistency was 44 per-
cent less likely to be found totally credible and 33 percent less likely to have
all of his testimony relied upon than was a witness with all the same charac-
teristics whose testimony did not contain a serious inconsistency. Although
these findings clearly suggest that Trial Chambers’ assessments of witness testimony are appropriately influenced by the presence of serious inconsistencies, the data also showed that for about one-third of witnesses with serious inconsistencies, the Trial Chambers nonetheless relied on the seriously inconsistent testimony to find the fact in question. In addition, for about one-quarter of the witnesses with serious inconsistencies, the Trial Chambers never even mentioned the inconsistencies in their judgments.

A decline in the Trial Chambers’ willingness to credit witnesses or rely on their testimony occurred across virtually all subpopulations of witnesses in the dataset. However, that decline was far more dramatic for some populations than others. These included male witnesses, Hutu witnesses, witnesses imprisoned for genocide, and witnesses with serious inconsistencies in their testimony. Finally, the data shows a decline over time in the Trial Chambers’ willingness to rely on seriously inconsistent testimony to find facts. Relatedly, the data shows an increase over time in the Trial Chambers’ willingness to mention serious inconsistencies in their judgments.

B. Understanding the Findings

International criminal law has its share of critics, and my previous fact-finding work placed me among them. However, the data resulting from this larger-scale and more rigorous empirical study tells a more nuanced and a more heartening story. The data depicts an international criminal justice system that, throughout its life, confronted a constant stream of serious testimonial deficiencies. There was once reason to believe that the incidence of these deficiencies would decline over time, and the fact that they did not provides clues as to their causes. What did decline, however, was the Trial Chambers’ willingness to credit prosecution witnesses and rely on their testimony. Indeed, the data arguably depicts an international criminal justice system that, over time, strengthened its commitment to factual accuracy and, more broadly, to the beyond-a-reasonable-doubt standard for convictions.

1. Understanding Serious Inconsistencies

Because serious inconsistencies stand as the most challenging testimonial deficiency in international criminal law—a realm featuring no shortage of challenging testimonial deficiencies—it is important to learn what we can about the factors that predict them. What my data suggests, however, is that not a lot of factors do predict them. Certainly, the one statistically significant finding—that witnesses who provide more pretrial statements/testimonies are more likely to testify seriously inconsistently—is one that common sense might lead us to expect. However, other expectations, which might seem equally grounded in common sense, were not supported by the data. For instance, although I was not surprised to learn that neither ethnic-
ality nor gender predicted serious inconsistencies, I did expect that imprisonment status and accomplice status would. My expectation stemmed from the fact that Rwandan courts and gacaca courts offer leniency to defendants who confess their crimes, but in order for a confession to “count,” as it were, the defendant must inculpate other offenders.\(^{131}\) Thus, because Rwandan and gacaca courts provide significant incentives for naming other offenders, many commentators believe that imprisoned defendants falsely accuse others of genocide.\(^{132}\) Assuming that is true, and reasoning that false testimony is more likely to contain serious inconsistencies than truthful testimony, I expected the rate of serious inconsistencies to be higher among witnesses imprisoned for genocide than among witnesses who were not. However, the logit regression did not support my expectation. One possible explanation for this finding is that imprisoned witnesses on average submitted nearly twice the number of statements/testimonies as non-imprisoned witnesses did.\(^{133}\) Because the number of documents submitted is a statistically significant predictor of serious inconsistencies, perhaps the effect of imprisonment status is felt indirectly through the number of documents submitted.

That said, even some of the tabular data regarding the relationship between serious inconsistencies and imprisonment status proved surprising. The tabular data did show that non-accomplice and non-imprisoned witnesses had lower rates of serious inconsistencies than accomplice and imprisoned witnesses, as expected, but it also showed that witnesses who were imprisoned at the time of trial had lower rates of serious inconsistencies than did witnesses who had been imprisoned but released before trial. Those statistics are counterintuitive if we assume that false testimony is more likely to contain serious inconsistencies. That assumption leads us to expect that a greater proportion of witnesses who were imprisoned at trial would testify falsely than would witnesses who were already released, because currently imprisoned witnesses would seek leniency through false accusations, whereas witnesses who had already been released would have nothing to gain.

Given that the data does not support my expectations, we need to unpack and more carefully scrutinize the assumptions underlying my expectations. For instance, I assumed that witnesses who were imprisoned at trial were


\(^{133}\) Witnesses who were not imprisoned provided an average of 1.44 statements per witness, whereas witnesses who were currently imprisoned or imprisoned but released provided averages of 3.44 and 3.7 statements/testimonies per witness, respectively.
more likely to testify falsely than other witnesses, but that may not be true. Certainly, some imprisoned witnesses had powerful incentives to falsely incriminate ICTR defendants, but some Tutsi victims did as well. Indeed, many of the most well known allegations of perjury at the ICTR have centered on Tutsi victims groups whose members allegedly encouraged other members to falsely accuse ICTR defendants. Moreover, even if it is true that imprisoned witnesses were more likely to testify falsely than non-imprisoned witnesses, that does not necessarily mean that their false testimony was more likely to contain serious inconsistencies. Although the presence of a serious inconsistency may be a predictor of false testimony, the reverse may not be true. Indeed, it is possible that perjuring witnesses—and particularly perjuring witnesses who are lying in order to receive tangible and substantial benefits—take greater care than truthful witnesses to keep their representations consistent. Finally, at least with respect to the tabular data, the counterintuitive statistics may stem from the small sample size of witnesses who were imprisoned but released before trial. Whereas the dataset included 264 witnesses who had never been imprisoned and fifty-two witnesses who were imprisoned at trial, it included only twenty witnesses who had been imprisoned but released before trial.

Although witness characteristics are not generally predictive of serious inconsistencies, certain characteristics of pre-trial statements/testimonies are. I explored three: (1) the number of statements/testimonies that a witness submitted; (2) the length of time between the statements/testimonies and the subject ICTR testimony; and (3) the type of statements/testimonies that a witness submitted. As for the number of witness statements/testimonies, the logit regression showed that the more statements/testimonies a witness provided, the greater the likelihood that the witness's subject testimony would be seriously inconsistent with at least one of the statements/testimonies. As for the length of time between pretrial statements/testimonies and subject testimony, the data shows no correlation. Indeed, tabular results suggest that ICTR testimony is just as likely to be seriously inconsistent with a statement that is more than thirteen years old as with a statement

---


135. Certain research suggests this possibility. See Vredeveldt, et al., supra note 47, at 184. See also Prosecutor v. Muvunyi, Case No. ICTR 00-55A-T, Judgment, para. 14 (Feb. 11, 2010) [hereinafter Muvunyi Judgment] (observing that "[w]hen an accomplice witness testifies in accordance with a prior statement implicating the accused, a Trial Chamber must be mindful that the witness may have had a motive or incentive to implicate the accused when he gave the prior statement, even if he has already been sentenced or has served his sentence").
that is less than one year old.\textsuperscript{136} Turning finally to the type of statements/testimonies that witnesses submitted, the data shows that Rwandan court statements/testimonies were most likely to give rise to a serious inconsistency, followed in descending order by 	extit{gacaca} statements/testimonies, ICTR statements, ICTR testimony, and finally foreign court statements/testimonies. In some cases, the differences were dramatic. For instance, nearly twice the percentage of Rwandan court statements/testimonies (42 percent) gave rise to serious inconsistencies as foreign court statements/testimonies (22 percent), and Rwandan court statements/testimonies were also considerably more likely to give rise to serious inconsistencies than ICTR pre-trial statements (31 percent).

At first glance, the latter two findings might appear counterintuitive. We might expect testimony to be more likely to be seriously inconsistent with older statements/testimonies than with newer ones because empirical studies of witness testimony show that memories of events fade over time.\textsuperscript{137} Moreover, we might expect ICTR statements to give rise to a larger percentage of serious inconsistencies than Rwandan court or 	extit{gacaca} statements/testimonies because the former are apt to contain more errors than the latter. ICTR investigators were almost certainly less informed than their Rwandan counterparts about the subjects of their investigations.\textsuperscript{138} Interviews conducted by John Jackson and Yassin Brunger, for instance, revealed that early ICTR statements, in particular, were apt to feature inaccuracies because investigators did not understand the information they were being provided.\textsuperscript{139} For example, Jackson and Brunger reported that ICTR investigators “often failed to be culturally sensitive . . . ‘blundering in’, as one respondent described it, by asking a series of inappropriate questions that could upset people.”\textsuperscript{140} Because Rwandan investigators understood the culture and the context, we might expect their statements to feature fewer inaccuracies. Another reason to expect fewer mistakes in Rwandan court and 	extit{gacaca} statements/testimonies is because they are taken in the witnesses’ native language; ICTR interviews, by contrast, went through at least two rounds of

\textsuperscript{136} See Table 8.
\textsuperscript{137} See John W. Shepherd et al., Identification Evidence: A Psychological Evaluation 80–86 (1982) (describing authors’ study, which showed that memory remained relatively constant for a few months after an event but declined sharply after eleven months); Hadyn D. Ellis, Practical Aspects of Face Memory, in Eyewitness Testimony: Psychological Perspectives 12, 23–25 (Gary L. Wells & Elizabeth F. Loftus eds., 1984) (summarizing research).
\textsuperscript{138} Nicola Palmer, Courts in Conflict: Interpreting the Layers of Justice in Post-Genocide Rwanda 70–71 (2015) (observing that “sets of [ICTR] investigators unfamiliar with the Rwandan context and heavily dependent on translation were operating in the harrowing reality of post-genocide Rwanda”).
\textsuperscript{140} Id. at 174.
language translations. Indeed, ICTR statements were drafted in English or French, so witnesses could not read the original versions before signing off on them.

These findings are counterintuitive, however, only if one assumes that serious inconsistencies result from innocent causes such as memory lapses or investigatory errors. Certainly, if a sizeable proportion of serious inconsistencies stemmed from witnesses' failure at the time of trial to remember the details that they knew at the time they gave their statements, then the incidence of serious inconsistencies would rise as the length of time between the statements/testimonies and the subject testimony grew. Similarly, if a sizeable proportion of serious inconsistencies was caused by investigators, who made mistakes when drafting witness statements, then ICTR statements would generate the largest proportion of serious inconsistencies, not Rwandan court statements/testimonies. To be sure, these innocent explanations can and almost certainly do cause some inconsistencies, but the findings described suggest that they do not drive a large percentage of them.

This is an important insight because the most pressing question concerning serious inconsistencies is not the who, the what, the where, or the when, but the why. Certainly, it is interesting to learn whether male witnesses testify seriously inconsistently at a greater rate than female witnesses, or whether early ICTR cases featured more serious inconsistencies than later cases. But what is truly valuable is to learn why witnesses testify seriously inconsistently. To be sure, serious inconsistencies present serious challenges to accurate fact-finding whatever their cause, but different causes give rise to different remedies and more importantly, different Trial Chamber responses. For instance, if we believe that most serious inconsistencies result from investigatory errors, as so many witnesses claim, then Trial Chambers should typically credit witnesses' in-court testimony, and prosecutors should take steps to improve the statement-taking process, perhaps by training investigators more thoroughly and ensuring better translation during the interviews. Those responses are not so appropriate, however, if we think that most inconsistencies arise because witnesses purposely tell different stories at different times.

141. Typically investigators ask questions in English or French, and the questions are interpreted into Kinyarwanda. The witness's answers—in Kinyarwanda—then must be interpreted back to English or French. See, e.g., Akayesu Judgment, supra note 15, at para. 145; Alexander Zahar, The ICTR's Media Judgement and the Reinvention of Direct and Public Incitement to Genocide, 16 CRIM. L. FORUM 33, 41 n.26 (2005).

Indeed, if we assume that the majority of serious inconsistencies is caused by witnesses purposely telling different stories at different times, then my findings are not counterintuitive at all. For instance, we might suppose that translation errors and investigator errors do cause some inconsistencies, but because those errors are apt to be somewhat random, the inconsistencies they cause are not so likely to be deemed "serious." In addition, a far greater proportion of witnesses who provided Rwandan court statements had been imprisoned for genocide than the witnesses who provided ICTR statements.\footnote{Twenty-one percent of witnesses who provided ICTR pre-trial statements had been imprisoned for genocide crimes, whereas 55 percent of witnesses who provided Rwandan courts statements/testimonies had.} As noted, a witness’s imprisonment status is not by itself predictive of serious inconsistencies, but combining that status with the criminal justice system in which the statement was taken could tell a different story. In particular, witnesses accused of genocide in Rwandan courts not only had incentives to falsely accuse ICTR defendants, but also to craft their stories so as to minimize their own responsibility. Trial Chambers regularly suspected imprisoned witnesses of engaging in such manipulations,\footnote{See, e.g., Kanyarukiga Judgment, supra note 66, at para. 578 & n.1607 (“Given that Witness CDL is still serving time in Rwanda for crimes related to the events of 1994, he could have personal motivations to implicate the Accused while minimizing his own role in the attacks.”); Ndahimana Judgment, supra note 58, at para. 244 (“As shown by the evidence in this case, the witness may have tried to minimize his own role in events.”); Setako Judgment, supra note 59, at para. 156 (“The potential, therefore, exists that the witness’s testimony may be influenced by a desire to positively impact his circumstances in Rwanda or to shift blame to Setako either to minimize his own involvement or based on the belief that Setako was behind his increased sentence.”); Munyakazi Judgment, supra note 55, at para. 119 (“[A]t earlier stages of the proceedings in Rwanda, the witness may have attempted to minimize his involvement in the genocide.”); id. at paras. 131, 309, 371, 420, 421; Prosecutor v. Muvunyi, Case No. ICTR 2000-55A-T, Judgment and Sentence, para. 156 (Sept. 12, 2006).} and if their suspicions were correct, the testimony of these witnesses could easily diverge from their previous representations.

Finally, one of the study’s most notable—and counterintuitive—findings is also consistent with the hypothesis that a large proportion of serious inconsistencies reflect purposefully false testimony. The notable finding is that the incidence of serious inconsistencies remained constant throughout the ICTR’s life. The finding is counterintuitive because, by all accounts, investigatory practices improved over time,\footnote{Report on the Conference on International Criminal Justice, Assembly of States Parties, ICC-ASP/6/INF-2, 42, 43 (Oct. 19, 2007); Jackson & Brunger, supra note 139, at 173–77. When the ICTY and ICTR were created, no international crimes had been investigated for fifty years. Thus, at that time, a cadre of competent professionals experienced in the investigation of mass atrocities did not exist. See Hassan B. Jallow, Challenges of International Criminal Justice: The ICTR Experience, OBSERVATIONS AT THE COLLOQUIUM OF PROSECUTORS OF INTERNATIONAL CRIMINAL TRIBUNALS (Nov. 25–27, 2004), http://ictr-archive09.library.cornell.edu/ENGLISH/colloquium04/jallow.html (noting that the ICTR “did not have the advantage of investigators with specialized capacity to investigate genocide, crimes against humanity or war crimes as no such tribunals previously existed”); Jackson & Brunger, supra note 139, at 169 (noting that when the ICTY and ICTR were created, “the United Nations was an organization with no previous experience in running criminal justice activities”). That cadre developed over the course of the ICTR’s life. Investigations practices probably improved, in addition, in response to the Trial Chambers’ increasing willingness to reject prosecution witness testimony that featured serious inconsistencies.} so those improvements would be
expected to produce a decrease in serious inconsistencies. But again, such a
decrease would appear only if a large proportion of serious inconsistencies
resulted from investigator errors. That the incidence of serious inconsisten-
cies remained constant is yet another indication that investigatory errors did
not play a primary causal role in serious inconsistencies.

It goes without saying, first, that there is no one set of causes for all
inconsistencies, and second, that we will be unable to determine the causes
for most inconsistencies at an individual level. But it is possible that some
causes explain a sizeable proportion of serious inconsistencies, and I believe
that my data provides support for the view—frequently advanced by defense
counsel—that witness mendacity is a primary driver of serious inconsisten-
cies. Indeed, the findings discussed in the next section suggest that the Trial
Chambers came to largely the same conclusion.

2. Understanding the Trial Chambers’ Assessments

Just as witness characteristics were not predictive of serious inconsisten-
cies, they likewise were not predictive of the Trial Chambers’ credibility and
reliability findings. Neither gender nor ethnicity was predictive, and neither
accomplice nor imprisonment status was predictive. The latter finding was
unexpected because the Trial Chambers repeated—in judgment after judg-
ment—their recognition that accomplice and imprisoned witnesses had in-
centives to lie, and their practice of viewing the testimony of accomplice
and imprisoned witnesses “with caution.” My findings might suggest,
then, that the Trial Chambers’ actions did not match their words. Indeed,
the testimony of accomplices was often crucial to the charges against the
defendant, so, to the extent that the Trial Chambers did not want to
acquit large numbers of defendants, they could have been motivated to rely
on the testimony of accomplice witnesses even though they ostensibly
viewed it with caution. At the same time, it is possible that the Trial Cham-
bbers did in fact view the testimony of accomplice and imprisoned witnesses

146. See, e.g., Prosecutor v. Nchamihigo, Case No. ICTR 01-63-T, Judgment and Sentence para. 17
(Nov. 12, 2008); Renzaho Judgment, supra note 59, para. 166; Muvunyi Judgment, supra note 135, para.
14. The Appeals Chamber repeatedly highlighted the same concern. See, e.g., Prosecutor v. Niyitegeka,
Case No. ICTR 96-14-A, Judgment, para. 98 (July 9, 2004) [hereinafter Niyitegeka Appeal Judgment];
Nchamihigo Appeal Judgment]; Prosecutor v. Kanyarukiga, Case No. ICTR 02-78-A, Judgment

147. For examples of judgments that use this phrase, see note 115.

148. See, e.g., Nzabonimana Judgment, supra note 58, at paras. 1224–1225 (relying on accomplice
witness CNAA to find that the defendant accused bourgmasters of not supporting the killings of Tutsis and
warned that they would be replaced by Interahamwe); Kanyarukiga Judgment, supra note 66, at para. 183
(relying on accomplice witness CBR); id. at para. 472 (relying on accomplice witness CBR to find that
defendant was with Ndashimana when he asked the attackers to “start working”); Renzaho Judgment,
supra note 59, at para. 251 (relying on accomplice witnesses AWE and GIJ to find that the defendant had
ordered individuals to collect and distribute weapons); id. at para. 652 (relying on accomplice witness
BUO to conclude that the defendant ordered the Interahamwe to attack, and later, to stop the killings at
Sainte Famille on June 17).
more skeptically, but that the testimony withstood the additional scrutiny. Accomplice witness testimony may have seemed particularly reliable, for instance, because accomplice witnesses often knew more than non-accomplice witnesses about the events in question, and specifically about the defendant's conduct. For this reason, their testimony may have been given greater weight. At the same time, accomplice witness testimony may have seemed particularly reliable because accomplices had the time to create that perception. For instance, some witnesses who were imprisoned together concluded to concoct false but consistent allegations.\textsuperscript{149} Trial Chambers frequently sought corroboration for accomplice witness testimony,\textsuperscript{150} but allegations generated through collusion did corroborate one another, and they thereby may have generated more positive credibility and reliability findings than one would expect after a cautious review of the testimony.

The study's two most notable findings are also the two most heartening. First, Trial Chambers were appreciably less likely to credit a witness or rely on his testimony if that testimony contained a serious inconsistency. That finding suggests that Chambers appropriately recognized that serious inconsistencies usually reflect unreliability and often mendacity. Second, the fact that Trial Chambers became less willing to find witnesses credible and to rely on their testimony as time passed reflects the same recognition.

When Trial Chambers manifest a statistically significant change in their treatment of testimony, two explanations are possible. First, the change might reflect a change in the quality of the testimony. Second, the change might reflect a change in the Trial Chambers' assessment methods.\textsuperscript{151} Turning to the first explanation, we can certainly point to changes that occurred over the course of the ICTR's life that could have reduced the quality of witness testimony. For example, my data shows that the ethnic composition of witnesses changed over time.\textsuperscript{152} In addition, in 2005, Rwanda began conducting \textit{gacaca} trials\textsuperscript{153} that often addressed the same crimes at issue in ICTR cases.\textsuperscript{154} Although these changes had the potential to reduce the quality of witness testimony, the data suggests that they did not. As already reported, witness characteristics such as ethnicity were not statistically significant predictors of serious inconsistencies or the Trial Chambers' testimo-


\textsuperscript{150} Prosecutor v. Muvunyi, Case No. ICTR 00-55A-A, Judgment, para. 131 (Aug. 29, 2008); Nchamihigo Appeals Judgment, supra note 146, para. 47.

\textsuperscript{151} It is also possible that both explanations combined to produce the change; however, in this case, this possibility is unlikely, as explained in the text.

\textsuperscript{152} The proportion of Tutsi witnesses declined by an average of approximately 2.2 percent per year.


\textsuperscript{154} Nicola Palmer reported that "since the nationwide implementation of \textit{gacaca} in 2005, forty-seven out of forty-nine ICTR cases have discussed evidence gathered by the \textit{gacaca} courts." Palmer, supra note 138, at 83.
nial assessments, so there is no reason to think that a change in the ethnic composition of witnesses would produce a change in the Trial Chambers' credibility or reliability findings. As for gacaca proceedings, their most likely effect would have been to increase the incidence of serious inconsistencies. To be sure, the introduction of gacaca almost certainly increased the average number of statements/testimonies that witnesses provided, and that factor did correlate with an increased incidence in serious inconsistencies. But the overall incidence of serious inconsistencies remained constant over time, as did the proportion of imprisoned witnesses. So we have no basis for believing that the introduction of gacaca led to a decline in the quality of witness testimony.

Finding no evidence of a decline in the quality of witness testimony, we must assume that testimonial quality remained relatively constant, but that the Trial Chambers' attitude toward that testimony changed over time. What might account for that attitude change? I believe it stemmed from a combination of two forces, one largely external and the other largely internal. First, the change reflects the Trial Chambers' increasing awareness of the problems associated with prosecution witness testimony. Second, the change reflects a maturation process that included the ICTR but extended to international criminal justice more broadly.

The first explanation posits that Trial Chambers became less willing to rely on prosecution witness testimony as time passed because, as time passed, Trial Chambers became increasingly aware that a non-trivial proportion of prosecution witness testimony contained inaccuracies as well as lies. During the earliest cases in my dataset, defense counsel were unable even to set foot in Rwanda, so their ability to collect evidence that would cast doubt on prosecution witness testimony was highly constrained. It should come as no surprise, then, that Trial Chambers found virtually every prosecution witness in those early cases credible. As time went on, however, defense counsel became better able to conduct investigations that called prosecution witness testimony into serious question. At this same time that ICTR defense investigations were producing more evidence of witness mendacity, the reports of genocide trials in Rwanda were telling a similar story. Scholars and human rights organizations covering the trials reported a high incidence of perjury, and the subsequent introduction of gacaca and the extraordinary attention it generated only increased awareness of false accu-

---

157. For a small sampling of the literature on gacaca, see, e.g., CLARK, supra note 77; PAUL CHRISTOPH BORNKAMM, RWANDA’S GACACA COURTS (2012); TERTSAKIAN, supra note 131; Waldorf, supra note 132; Sosnov, supra note 132.
sations. The very fact that Rwandan courts, which were widely believed to manifest a pro-conviction bias, acquitted more than 25 percent of defendants was itself compelling evidence that many Rwandans falsely accused their compatriots. Finally, the publication of my book Fact-Finding Without Facts also publicized testimonial deficiencies in general and false testimony in particular.

Therefore, as time passed, Trial Chambers could not help but become aware that at least some prosecution witnesses testified falsely. Awareness does not always lead to action, but in this case, a variety of forces combined to motivate Trial Chambers to alter their treatment of witness testimony in light of their newfound awareness. One source of particularly compelling motivation derived from the ICTR Appeals Chamber, which occasionally overturned Trial Chamber factual findings that it considered insufficiently supported or insufficiently defended, and which issued judgments that sought to guide Trial Chamber assessments of evidence. For instance, the Appeals Chamber delineated factors that Trial Chambers should consider when assessing accomplice witness testimony and testimony containing serious inconsistencies. The Appeals Chamber also admonished the Trial Chambers to describe their consideration. As the Appeals Chamber stated in Renzaho, "Trial Chambers cannot merely state that they exercised caution when assessing the evidence of an accomplice witness, but must establish that they in fact did so."

These admonitions, and in particular their focus on accused and imprisoned witnesses and witnesses with serious inconsistencies, may help to explain the more granular time trends revealed by my data. Recall that, over time, Trial Chambers became less willing to rely on the testimony of all subcategories of prosecution witnesses, but the trend was far more dramatic for accomplice and imprisoned witnesses and witnesses with serious inconsistencies than for other groups. These trends are in keeping with the Ap-

---

158. See supra note 132.
160. Id. at 59.
161. FACT-FINDING WITHOUT FACTS, supra note 7, at 130–66.
165. Id. at para. 47.
peals Chamber's dicta as well as with the Trial Chambers' increasing awareness of the prevalence of false testimony.

Finally, Finding Without Facts may also have played a role in motivating Trial Chambers to take testimonial deficiencies more seriously. It was the first publication to provide an in-depth look at the problematic features of international criminal testimony and, more importantly, at the Trial Chamber's treatment of that testimony. By criticizing the Trial Chambers for adopting "a cavalier attitude toward testimonial deficiencies" and for frequently basing "their convictions on deeply flawed testimony," my book may have hit a nerve and encouraged greater scrutiny of prosecution witness testimony in the future.

Although I have little doubt that the ICTR's fact-finding evolution was driven in large measure by the facts on the ground that I have just described, I likewise have little doubt that the evolution also reflects the natural, and promising, maturation process of international criminal justice in general and the ICTR in particular. The ICTR and ICTY began their lives as weak and vulnerable institutions whose long-term survival was in grave doubt. In those days, international criminal justice was an experiment that many expected to fail. And no wonder. Initially, the Tribunals were provided inadequate funding, inadequate enforcement tools, and inadequate political support. The ICTR was particularly dependent on Rwanda for access to crime sites and witnesses, so it generally took care to stay in the country's good graces. At the same time, the Tribunals bore all

---

168. Fact-Finding Without Facts, supra note 7, at 189. In addition, I argued, "Trial Chambers routinely discount the significance and potentially distortive effect of many testimonial deficiencies." Id. at 199.

169. For an empirically grounded contention that the ICTR has sought to legitimate its work, see Palmer, supra note 138, at 160-63.


175. See Victor Peskin, International Justice in Rwanda and the Balkans 151-231; Cruvellier, supra note 30, at 10-13; Jackson & Brunger, supra note 139, at 170. The ICTR's high level of dependence on Rwanda became especially clear in 1999, after the Appeals Chamber dismissed with prejudice the indictment of Jean-Bosco Barayagwiza and ordered him released after concluding that the prosecution had violated Barayagwiza's right to be brought promptly before a judge following arrest. Prosecutor v. Barayagwiza, Case No. ICTR 97-17-AR72, Decision (Nov. 3, 1999). The Appeals Chamber's decision outraged Rwanda to such a degree that it suspended all dealings with the ICTR and refused to issue a visa to the Tribunal's Chief Prosecutor, Carla Del Ponte. Peskin, supra at 177-85; Franck Petir, Cameroonian Intrigues, Int'l Just. Trib., Mar. 5, 2001.Del Ponte asked the Appeals Chamber to reconsider its decision, and she candidly acknowledged Rwanda's power over the Tribunal, when
the hopes and dreams of human rights advocates who saw international criminal justice as the new frontier in preventing mass atrocities.

I have argued previously that the vulnerability that characterized the Tribunals' early years influenced many of its early decisions, including its procedural choices and defense counsel regulations. I suspect that early vulnerability also played a role in the Trial Chambers' kindly treatment of prosecution witness testimony. With significant hurdles to overcome and even more significant expectations to fulfill, the early ICTR could not afford to reject large quantities of prosecution witness testimony. Elsewhere, I have explored the high costs of international criminal acquittals, but during the ICTR's early years, even the rejection of a substantial quantity of prosecution witness testimony would have proven extraordinarily costly. When established courts reject large quantities of prosecution witness testimony, the courts are lauded for their judicial independence. Had the early ICTY and ICTR rejected large quantities of prosecution witness testimony, they would have provoked a storm of criticism. It was bad enough that both Tribunals were delivering slow and expensive justice. At least they were delivering justice. If, following a genocide that killed more than half a million people, prosecutors showed themselves unable to collect evidence of a sufficient quality to withstand Trial Chamber scrutiny, then commentators likely would have called for the experiment in international criminal justice to come to an end.

Not surprisingly, then, ICTR Trial Chambers' scrutiny became easy to withstand. Commentators have observed that early ICTY and ICTR cases

she told the Chamber: "If I don't get cooperation from Rwanda, . . . I can first open the door at the detention center and set them all free and then second I can close the door to my office because without them I cannot do anything all." J. Coll Metcalf, An Interview with United Nations' Chief War Crimes Prosecutor, Carla del Ponte, INTERNEWS, Feb. 15, 2000. In response, the Appeals Chamber reinstated Barayagwiza's indictment, and relations between the ICTR and Rwanda returned to normal. Prosecutor v. Barayagwiza, Case No. ICTR 97-17-AR72, Decision on Prosecutor's Request for Review or Reconsideration (Mar. 31, 2000). For an insider's account of the case, see MOGHALU, supra note 10, at 101-23.

176. The Tribunals were considered the testing ground for a permanent international criminal court, then under negotiation, see Geoffrey R. Watson, The Changing Jurisprudence of the International Criminal Tribunal for the Former Yugoslavia, 37 NEW ENG. L. REV. 871, 872 (2003), and the promise of international justice was proclaimed by scholars and practitioners alike, see, e.g., Richard Goldstone, Conference Luncheon Address, 7 TRANSNAT'L L. & CONTEMP. PROBS. 1, 2 (1997); Richard May & Marieke Wierda, Evidence Before the ICTY, in ESSAYS ON ICTY PROCEDURE AND EVIDENCE IN HONOUR OF GABRIELLE KIRK MCDONALD 249, 252-53 (Richard May et al. eds., 2001); RUTI G. TEITEL, TRANSITIONAL JUSTICE 56 (2000); Antonio Cassese, On the Current Trends Towards Criminal Prosecution and Punishment of Breaches of International Humanitarian Law, 9 EUR. J. INT'L L. 2, 9-10 (1998); Antonio Cassese, Reflections on International Criminal Justice, 61 MOD. L. REV. 1, 6-9 (1998); Stephan Landsman, Alternative Responses to Serious Human Rights Abuses: Of Prosecution and Truth Commissions, 59 LAW & CONTEMP. PROBS. 81, 83 (1996).

177. Combs, supra note 6, at 377-80.

178. See FACT-FINDING WITHOUT FACTS, supra note 7, at 228-34.

179. See, e.g., John E. Ackerman, Assignment of Defence Counsel at the ICTY, in ESSAYS ON ICTY PROCEDURE AND EVIDENCE IN HONOUR OF GABRIELLE KIRK MCDONALD 167, 170 (Richard May et al. eds., 2001) ("One of the major criticisms levelled at the Tribunal is the length of trials."); Combs, supra note 173, at 90-94.
often were "thinly investigated" and likely to be based on "weak evidence." Professor and former prosecutor Alex Whiting has described the way in which ICTY judges, who joined with prosecutors in desiring the tribunals' success, provided prosecutors tremendous latitude to amend charges and add evidence long after a defendant's arrest. ICTR judges were likewise lenient in that regard, and, as this study shows, in their assessments of witness testimony. In early cases, deficiencies in prosecution witness testimony were largely ignored.

In the intervening years, international criminal justice came a long way. The ICTY and ICTR not only carried out their own mandates, but they also paved the way for a host of new ad hoc tribunals as well as for a permanent international criminal court. And as the tribunals themselves grew—from novel and vulnerable institutions to the foundational bodies of the global international criminal justice system—something very positive happened to fact-finding at the ICTR. No longer confronting the threat of closure, ICTR judges began to more carefully scrutinize prosecution witness testimony, highlight its deficiencies, and reject substantial quantities of it. The judges' increased awareness of false testimony combined with the enhanced credibility and legitimacy of their tribunal to enable ICTR judges to engage in more scrupulous, defensible fact-finding.

VI. CONCLUSION

As Rosemary Byrne has observed, "Empirical research on international criminal trials can lead to a recalibrating of assumptions about the international criminal trial process, and in turn offer an informed basis upon which to improve international trial practice." The empirical study described here, though centering on the ICTR, has broad implications for international criminal justice as a whole. For one thing, the current international tribunals struggle with the same fact-finding impediments that made truth so elusive at the ICTR. In particular, serious inconsistencies appear with

180. Whiting, supra note 13, at 137. Former ICTR Special Counsel and Spokesman Kingsley Moghalu diplomatically conceded that the ICTR's prosecution "was frequently not 'trial-ready' before obtaining arrest warrants ... " Moghalu, supra note 10, at 193; see also Jallow, supra note 145 (noting that "in the early years arrests and detention of suspects frequently preceded investigations").


182. Whiting, supra note 13, at 138.

183. See, e.g., Prosecutor v. Kabiligi & Ntabakuze, Case Nos. ICTR-97-34-I & ICTR-97-30-I, Decision on the Prosecutor's Motion to Amend the Indictment (Oct. 8, 1999); Prosecutor v. Simba, Case No. ICTR-01-76-I, Decision on Motion to Amend Indictment (Jan. 26, 2004); Prosecutor v. Zigiranyirazo, Case No. ICTR-01-73-I, Decision on Prosecutor's Request for Leave to Amend Indictment and on Defence Urgent Motion for an Order to Disclose Supporting Material in Respect of the Prosecutor's Motion for Leave to Amend the Indictment (Oct. 15, 2003).

184. See PALMER, supra note 138, at 60-61.

alarming frequency in the transcripts of ICC\textsuperscript{186} and the Extraordinary Chambers in the Courts of Cambodia.\textsuperscript{187} Moreover, the fact that the incidence of serious inconsistencies remained constant throughout the ICTR’s life, despite an increasing budget and a decreasing willingness on the part of Trial Chambers to rely on such testimony, should motivate prosecutors at the current tribunals to redouble their efforts to improve the quality of prosecution witness testimony. Even simple reforms to statement-taking could have considerable impact. For instance, at present, many witness statements are summaries of the information that the witness conveyed to the investigator. Consequently, when a statement fails to contain an allegation that is central to the witness’s testimony, the witness can plausibly, but perhaps inaccurately, explain the omission by claiming that he was never asked about the allegation. At the very least, then, witness statements should be drafted in transcript format so as to include both the investigator’s questions as well as the witness’s responses. The Niyitegeka Appeals Chamber recommended this very reform back in 2004,\textsuperscript{188} though it never became a consistent practice at the ICTR. Indeed, this is the most minor of the measures that should be


\textsuperscript{188} The Appeals Chamber stated that an “ideal” record of a witness interview would include “all the questions that were put to a witness and all the answers given by the witness” along with “[t]he time of the beginning and the end of an interview, specific events such as requests for breaks, offering and accepting of cigarettes, coffee, and other events that could have an impact on the statement or its assessment.” Niyitegeka Appeal Judgment, supra note 146, at para. 31. The Appeals Chamber concluded that it was “necessary to disclose the questions put to the witness in order to make the statement intelligible” and to enable the defense “to prepare for cross-examination properly.” \textit{Id.} at para. 33. Indeed, the Chamber went so far as to maintain that “it may be impossible to assess the probative value of the witness’s answer without juxtaposing it with the relevant question.” \textit{Id.}
The ICTR’s inability to reduce serious inconsistencies stands as a warning to today’s international tribunals to take testimonial deficiencies seriously.

The warning should be particularly ominous for the ICC because that court faces far more profound fact-finding challenges than those the ICTR confronted. The ICTR had twenty years to investigate one set of crimes in one small nation that occurred over the course of 100 days. The ICC, by contrast, must simultaneously investigate a series of crimes committed as part of a series of long-term conflicts occurring in locations all over the globe. ICC investigators will never have the luxury of familiarizing themselves with their subjects to the degree or depth that ICTR investigators were able to. Further, the conditions greeting ICTR investigators were far more secure and hospitable than ICC investigators can hope to enjoy. By the time the ICTR was created, the conflict in Rwanda had ended, so ICTR investigators could conduct their investigations without fear of violence. To be sure, Rwanda and the ICTR have had their disagreements over the years, but as a general matter, Rwanda approved of the ICTR’s prosecutions, and welcomed ICTR investigators into the country. By contrast, ICC investigators are unable even to enter some of the states where they should be investigating crimes, and when they are able to conduct in-state investigations, they, and prospective witnesses, frequently face such serious security risks that they must work through intermediaries and rely excessively on reports drafted by the United Nations or non-governmental organizations. Finally, although ICTR prosecutors regularly presented witnesses whose testimonies diverged from their previous representations, at least ICTR prosecutors could have confidence that their witnesses would show up and would inculpate the defendants. Many ICC witnesses, by contrast, have been intimidated into recanting their allegations and refusing to testify. Indeed, the prosecution’s failure to submit credible and reliable

189. For other suggested reforms, see FACT-FINDING WITHOUT FACTS, supra note 7, at 274–321.
190. The Rwandan conflict ended in July 1994, see 1 VIRGINIA MORRIS & MICHAEL SCHARF, THE INTERNATIONAL CRIMINAL TRIBUNAL FOR RWANDA 58 (1997), and the ICTR was created in November 1994, see id. at 72.
191. As a member of the Security Council, Rwanda voted against the creation of the ICTR, see id. at 72, and it completely obstructed the Tribunal’s efforts to prosecute members of the ruling Rwandan Patriotic Front, see Victor Peskin, Victor’s Justice Revisited: Rwandan Patriotic Front Crimes and the Prosecutorial Endgame at the ICTR, in REMAKING RWANDA: STATE BUILDING AND HUMAN RIGHTS AFTER MASS VIOLENCE 173, 173 (Scott Strauss & Lars Waldorf eds., 2011). Rwanda also believed the Tribunal cost too much for the number of offenders it prosecuted. CLARK, supra note 77, at 136.
194. Id. at 70.
195. See id. at 54–56.
196. See, e.g., Public Redacted Version of “Prosecution’s Request for the Admission of Prior Recorded Testimony of [REDACTED] witnesses” para. 2, Prosecutor v. Ruto & Sang (May 21, 2015) (Case No. ICC-01/09-01/11); Tom Maliti, Prosecutor Withdraws Seven Witnesses in Kenyatta Case in Past Year, INT’L
evidence has already emerged as a crucial early challenge for the ICC. ICC prosecutors have been the target of blistering criticism, most notably from ICC judges, but also from commentators who have questioned the prosecution’s “small team” approach to investigations, criticized its failure to conduct more in-state investigations, and urged it to adopt a more informed approach to investigations. This study, then, provides additional and concrete support for those views. The very fact that serious inconsistencies appeared in the testimony of approximately one-half of ICTR prosecution witnesses—throughout the ICTR’s life, when ICTR prosecutors labored under dramatically more favorable conditions—should serve as a wake-up call to the ICC.

And a loud wake-up call it should be. ICC judges have already shown unparalleled willingness to reject the prosecution’s evidentiary offerings. Unlike ICTR judges who initially called most doubts in favor of prosecution witnesses, ICC judges have, from the court’s very inception, required prosecutors to meet stringent evidentiary standards. In addition, they have backed up those standards with real consequences. Not only have ICC judges rejected large quantities of prosecution evidence in the two cases that resulted in conviction, they have refused to confirm charges in approximately one-third of the prosecution’s cases, and they have acquitted one out of three defendants. Some have suggested that the judges’ evidentiary requirements might be too exacting, but what is clear is that ICC prosecutors have no choice but to devote considerable attention to the sorts of testimonial deficiencies that ICTR prosecutors initially had the luxury to ignore.

In fact, whereas the ICC prosecution should take to heart this study’s findings on testimonial deficiencies, the ICC judges need no lessons from the ICTR on the treatment of witness testimony. I have argued that international criminal justice matured tremendously during the years the ICTR conducted trials, and the careful judicial scrutiny now on display at the ICC

197. See, e.g., Ngudjolo Judgment, supra note 12, at paras. 115–123.
198. See WAR CRIMES RESEARCH OFFICE, AM. UNIV. WASH. COLL. OF LAW, INVESTIGATIVE MANAGEMENT, STRATEGIES, AND TECHNIQUES OF THE INTERNATIONAL CRIMINAL COURT'S OFFICE OF THE PROSECUTOR 4–5, 24–30 (2012) (hereinafter INVESTIGATIVE MANAGEMENT). But see Alex Whiting, Dynamic Investigative Practice at the International Criminal Court, 76 L. & CONTEMP. PROBS. 163, 175 (2014) (arguing that the War Crimes Research Office "has the cause and effect precisely backwards").
201. Whiting, supra note 13, at 139–40.
202. INVESTIGATIVE MANAGEMENT, supra note 198, at 15.
stands as a compelling testament to the newfound legitimacy and credibility of the current international criminal justice system. As Julie O'Sullivan wrote about the ICC, "[f]ar from being a kangaroo court where verdicts of 'guilty' can be counted upon, the court is an entity where the prosecution is viewed with a skeptical and sometimes even jaundiced eye." Though difficult for the prosecution in the short run, O'Sullivan maintains that "[a]ny augmentation of the Court's perceived impartiality has the power to enhance the credibility of each prosecutorial victory." 

Back in 1994, the ICTY asserted that, "the success of the Tribunal as a whole depends very much on the caliber of [the Prosecution's] investigative staff." That statement was not really true for the ICTR in its early years, but it became true. By the end of its life, the ICTR's fact-finding had undergone a dramatic evolution, an evolution whose legacy the ICC is carrying on.

206. Id. at 157.