

Grave Crimes and Weak Evidence: Fact-Finding Evolution in International Criminal Law *PART III-IV*

Nancy Amoury Combs^{a, *}

a. William & Mary Law School, Williamsburg, Virginia, U.S.

*CORRESPONDING AUTHOR. E-mail: ncombs@wm.edu

ABSTRACT 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.

KEY WORDS Forensic identification, Craniofacial superimposition, Skull-face overlay, Coevolution, Cooperative coevolutionary algorithm, Fuzzy landmarks, Evolutionary algorithms, Forensic science

III. SERIOUS INCONSISTENCIES: THE WHO, WHAT, WHERE, AND WHEN OF TESTIMONY THAT DIVERGES FROM PREVIOUS REPRESENTATIONS

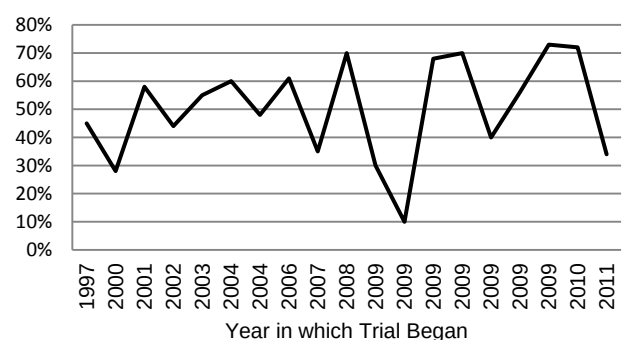
A. The Incidence of Serious Inconsistencies

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 % of witnesses in the dataset testified in a way that was inconsistent to some degree with the witnesses' previous statements/testimonies. When we consider only inconsistencies that are serious, the %age declines, though not dramatically so. The available transcripts and judgments show that 42 % 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,^[88] the %age of witnesses who testified

seriously inconsistently rose to 48 %.

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.^[89]



Graph 1: Showing chronological incidence of serious inconsistencies

⁸⁷ See FACT-FINDING WITHOUT FACTS, *supra* note 7, at 106-18.

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.

C. Serious Inconsistencies by Gender and Ethnicity

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 % for women and 44 % for men). There is a slightly greater difference in the incidence of serious inconsistencies between the two ethnic groups (41 % for Tutsi and 48 % 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 % for male Hutu compared to 44 % for male Tutsi), a far greater proportion of female Hutu witnesses testified seriously inconsistently (71 %) than female Tutsi witnesses (36 %).

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, ^[91] 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, ^[92] 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 % of accomplices testified seriously inconsistently, only 40 % of non-accomplices did so. However, the logit regression, ^[93] 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 % of witnesses who were never imprisoned testified seriously inconsistently whereas 52 % of witnesses imprisoned during trial did. Moreover, a whopping 75 % of witnesses who were imprisoned but released before trial testified seriously inconsistently. 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 %age 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

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

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 particular result, but the logit regression shows a highly statistically significant effect for the number of documents. In

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.

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 Logistic 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.

⁹⁴ See Logit Regression 1: Explaining Serious Inconsistencies by Comprehensive Set of Explanatory Variables, *infra* app. 1.

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 %. Adding more statements/testimonies continues to increase the probability of serious inconsistencies, but to a lesser degree as the number of statements/testimonies increases.^[95]

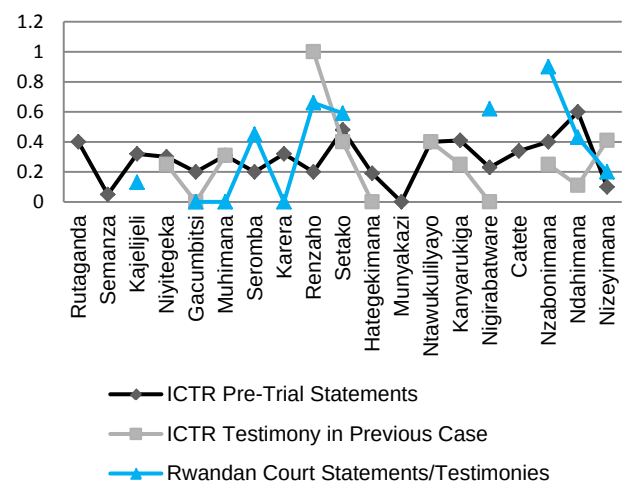
Overall, 33 % of statements/testimonies gave rise to serious inconsistencies, but these were not evenly distributed across the different kinds of statements/testimonies. For instance, statements/testimonies submitted to foreign courts generated the smallest proportion of serious inconsistencies whereas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

Table 7: Serious inconsistencies by document type

Type of Document	Percentage of those Documents Found Seriously Inconsistent with ICTR Type of Document Testimony in the Subject Trial
Rwandan Court Statements/Testimonies	42%
Gacaca Statements/Testimonies	35%
ICTR Pre-Trial Statements	31%
Testimony in Previous ICTR Cases	30%
Foreign court Statements/Testimonies ^[96]	22%
All Documents Combined	33%

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 testimony that was seriously inconsistent with ICTR pre-trial statements.^[97] 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 %age of serious inconsistencies in previous ICTR testimony decreased by an average of 2.6 % per year while the %age of serious inconsistencies in Rwandan court statements/testimonies increased by an average of 4 % per year.^[98] 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^[99]

F. Serious Inconsistencies by Length of Time between Statements/Testimonies and the Subject Testimony

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 inconsistently 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.

IV. ASSESSING TESTIMONY: A COMPREHENSIVE EXAMINATION OF THE ICTR's CREDIBILITY AND RELIABILITY DETERMINATIONS

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.

⁹⁵ 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.

⁹⁶ The dataset included only nine pre-trial statements/testimonies from foreign courts, so this finding is less precise than the others.

⁹⁷ See Simple Regression 1: Serious Inconsistencies Over Time, by Document

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 %ages 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.

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.

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 % per year,^[100] and their positive total or partial credibility findings declined on average by 2.0 % per year.^[101] 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 % per year,^[102] whereas their willingness to totally or partially rely on a prosecution witness's testimony declined by an average of 2.1 % per year.^[103] 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 % the likelihood that a Trial Chamber would find a witness totally credible,^[104] and it reduced by 46 % the likelihood that a Trial Chamber would find a witness totally or partially

credible.^[105] Put another way, during the first trial in the dataset, a witness, who was average in every other respect, had a 97 % chance of being found totally credible and a 99 % chance of being found totally or partially credible. By the last trial in the dataset, that same witness had only a 46 % chance of being found totally credible and a 53 % 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 %^[106] and to wholly or partially rely on the witness's testimony by 51 %.^[107] That is, during the first trial in the dataset, a witness, who was average in every other respect, had an 81 % chance of having her testimony wholly relied upon, and a 97 % chance of having her testimony totally or partially relied upon. By the last trial in the dataset, that same witness had only a 23 % chance of having her testimony totally relied upon and a 46 % chance of having her testimony totally or partially relied upon.

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.

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.

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

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.

¹⁰⁴ See Logit Regression 2: Trial Chambers' Total Credibility Findings by Comprehensive Set of Explanatory Variables, *infra* app. 3.

¹⁰⁵ See Logit Regression 3: Trial Chambers' Total and Partial Credibility Findings by Comprehensive Set of Explanatory Variables, *infra* app. 3.

¹⁰⁶ See Logit Regression 4: Trial Chambers' Total Reliance on Testimony by Comprehensive Set of Explanatory Variables, *infra* apps. 3-4.

¹⁰⁷ See Logit Regression 5: Trial Chambers' Total and Partial Reliance on Testimony by Comprehensive Set of Explanatory Variables, *infra* app. 5.

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]

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.

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 % per year; it stayed virtually the same for female witnesses.^[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 % per year.^[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 % and 1.3 % per year, respectively.^[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 % and 4.5 % per year, respectively, for Hutu witnesses whereas they declined by an average of only 1.1 % and 1 % per year, respectively, for Tutsi witnesses.^[112] 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.^[113]

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 inculcate defendants,^[114] and they frequently claim to treat such witnesses' testimony "with caution."^[115] 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

¹¹¹ Summary 1, *supra* note 109; Summary 2, *supra* note 110.

¹¹² Summary 3: Simple Regression Statistics for Credibility Findings Over Time by Ethnicity, *infra* app. 6.

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.*

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.

¹¹⁵ See, e.g., Gacumbitsi Judgment, *supra* note 15, at para. 86; Gatete Judgment, *supra* note 58, at para. 405; Hategekimana Judgment, *supra* note 59, at paras. 278, 449, 547, 552; Kanyarukiga Judgment, *supra* note 66, at paras. 181, 198, 306, 440-441, 452-453, 487, 576, 591; Karera Judgment, *supra* note 15, at paras. 52, 165, 189, 215; Munyakazi Judgment, *supra* note 55, at paras. 10, 119, 199, 206, 255, 366, 417; Ndahimana Judgment, *supra* note 58, at paras. 48-49, 248-249, 443, 454, 459-61, 631, 687; Ngirabatware Judgment, *supra* note 57, at paras. 66, 193, 201, 283, 311, 479; Nizeyimana Judgment, *supra* note 59, at paras. 113, 504, 560, 608, 621, 811, 820, 836-838, 1107; Ntawukulilyayo Judgment, *supra* note 58, at paras. 199, 219, 233, 266, 434; Nzabonimana Judgment, *supra* note 58, at paras. 226, 1064, 1142, 1210, 1276, 1348, 1447, 1480; Renzaho Judgment, *supra* note 59, at paras. 166, 240, 312, 321, 410, 487, 557, 569, 594, 652, 734; Setako Judgment, *supra* note 59, at paras. 50, 72, 167, 200, 203, 250, 264, 339, 348, 367, 424.

¹⁰⁸ See Logit Regressions 2, 3, 4, and 5, *infra* apps. 3-5.

¹⁰⁹ Summary 1: Simple Regression Statistics for Credibility Findings Over Time by Gender, *infra* app. 6.

¹¹⁰ *Id.*; Summary 2: Simple Regression Statistics for Reliance Findings Over Time by Gender, *infra* app. 6.

relied on the testimony of witnesses who have never been imprisoned at approximately twice the rate that they did for imprisoned witnesses, though the divergence narrowed when witnesses whom the Trial Chambers found partially credible or whose testimony they relied on in part were included.

The tabular data for the Trial Chambers' treatment of accomplice witnesses shows a similar pattern. This is not surprising, as the population of witnesses who are accomplices overlaps substantially with the population of 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.

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]

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.

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 much.^[117] 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 counter-intuitively 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.

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 factfinding 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.

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 % less likely to be found totally credible^[118] and 35 % less likely to be found totally or partially credible than witnesses who do not testify seriously inconsistently.^[119] Similarly, Trial Chambers were 33 % less likely to totally rely on the testimony of a witness with a serious inconsistency^[120] and 35 % less likely to totally or partially rely on

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.

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.

¹¹⁶ See Logit Regressions 2, 3, 4, and 5, *infra* apps. 3-5.

the testimony of a witness with a serious inconsistency.^[121]

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

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 % per year.^[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 % 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 part on witnesses with serious inconsistencies declined at a much greater rate, namely an average of 3.4 % per year compared with an average of only 0.6 % 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 %age 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 %.

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 % per year.^[124]

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,^[125] they do not always mention serious inconsistencies.^[126] In particular, my dataset reveals that Trial Chambers mention about 75 % of serious inconsistencies in their judgments.^[127] 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 % per year.^[128]

V. SUMMARIZING AND EXPLAINING THE FINDINGS

(To Be Continued)

¹²¹ See Marginal Effects of Serious Inconsistencies on the Trial Chambers' Total and Partial Reliance Findings, *infra* app. 5.

¹²² Simple Regression Statistics for Credibility and Reliance Findings Over Time by Serious Inconsistencies in Testimony, *infra* app. 6.

¹²³ *Id.*

¹²⁴ Simple Regression 6: Trial Chamber's Reliance on Testimony with Serious Inconsistency to Find Facts, *infra* app. 2.

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.