Sanctioning Perpetrators of International Crimes: A Vignette Study

Catrien Bijleveld | ORCID: 0000-0002-2455-533X
Vrije Universiteit Amsterdam, De Boelelaan 1077, 1081 HV Amsterdam, The Netherlands
Corresponding author, e-mail: cbijleveld@nsr.nl

Margareta Blažević | ORCID: 0000-0002-0724-4492
Vrije Universiteit Amsterdam, De Boelelaan 1077, 1081 HV Amsterdam, The Netherlands
margareta.blazevic@gmail.com

Diana Bociga Gelvez | ORCID: 0000-0001-5751-3112
Vrije Universiteit Amsterdam, De Boelelaan 1077, 1081 HV Amsterdam, The Netherlands
diana.bociga@hotmail.com

Mirza Buljubasic | ORCID: 0000-0003-4169-5606
Faculty of Criminalistics, Criminology and Security Studies, University of Sarajevo, Zmaja od Bosne 8, Sarajevo 71000, Bosnia and Herzegovina
mirzabuljubasic@fkn.unsa.ba

Abstract

Limited research has been devoted to factors impacting the perceived justness of sentences for international crimes. We presented respondents with a vignette in which such a hypothetical crime was described, as well as some contextual information; in the vignette, the perpetrator received a sentence that was based on similar historical cases. In the study, the rank of the perpetrator, the apology by the perpetrator, and the location of the trial (in the country where the crime was committed or at the International Criminal Court) were manipulated. Respondents...
were asked whether they believed the sentence was just or whether they would hand down a different sentence. Qualitative questions were included to elucidate respondents’ choices. Our mixed methods analyses reveal how apology, over and above strong geographic differences, plays a significant and dominant role in perceived justness.

Keywords

atrocity crimes – apology – punishment – vignette design

1 Introduction

In 1996, Dražen Erdemović, a foot soldier, a member of the Bosnian Serb forces death squadron and one of the perpetrators of the Srebrenica genocide, stood before the newly established UN ad-hoc International Criminal Tribunal for the former Yugoslavia (ICTY). Before his final sentence of five years’ imprisonment was handed down, Erdemović had said, crying: ‘I feel sorry for all the victims, not only for the ones who were killed then at that farm, I feel sorry for all the victims’. Erdemović’s apology and demonstrated remorse, which were recognized as sincere by the judges, was one of the mitigating circumstances considered when determining his sentence.

Determining a just and appropriate punishment for a perpetrator of an international crime is a complex and challenging process. International crimes (e.g., war crimes, crimes against humanity and genocide) are a violation of peremptory norms (lat. ius cogens). The ramifications of such crimes are serious, destructive and long-lasting—both for the direct victims of the crime, their families and for affected society at large. Trials for international crimes

---

1 ICTY, Prosecutor v. Dražen Erdemović, Case No. IT-96-22, Trial Chamber, Sentencing Judgment, 29 November 1996.
3 Throughout the text, we will use the terms international crimes and atrocity crimes interchangeably to describe violations of international criminal law stricto sensu, that is, war crimes, crimes against humanity, genocide, and crime of aggression.
5 Ibid.
aim to determine facts about the events which unfolded, establish perpetrator accountability, mete out punishment and deliver recognition and acknowledgement of the victims' suffering.\(^6\) International criminal judges render sentences after close consideration of legal norms and procedures, mitigating and/or aggravating circumstances, victims’ suffering, various (e.g., contextual, structural and situational) factors, perpetrator role and actions, among others.\(^7\)

However, while international crimes are part and parcel of the singular body of international criminal law, the perceived fairness of sentences by the local (affected) population might be embedded in local punitive tradition or culture. To give an example, what is perceived as a just sentence for, say, rape in one country (on average two years’ imprisonment in the Netherlands)\(^8\) may be very different in another country (on average eight years’ imprisonment in England and Wales).\(^9\) The differences in this example are quite large in spite of these two countries being geographically close, and part of western Europe. An additional complication may be introduced by the fact that the public’s perceived fairness of sentencing is affected, as nascent literature indicates, not only by the nature of the sentence itself but also by the type of court that imposed it.\(^10\)

Judges’ perception or the law’s prescription of a fair and just sentence may all in all differ from what victims, (indirectly) affected population or the general public perceive as appropriate. Judicial proceedings and punishment of


perpetrators responsible for international crimes can influence the development of the official political narrative that can, in turn, influence attitudes towards punishment and courts, post-conflict social unity, and may encourage or discourage reconciliation. Therefore, perceptions of the appropriateness of punishment by people outside of the courtroom—victims, their families, affected populations or communities and the general public—are of paramount importance. What would the public’s opinion be regarding Erdemović’s judgement? What factors play a role in popular perceptions of whether judgments are fair and justice has been done? What if the Erdemović case was different (e.g., what if he had been tried at a national court, what if he had not been only a foot soldier, and what if he had not shown remorse)? To what extent, then, would lay opinion differ?

Empirical research on the perceived fairness of international crimes sentencing has been scarce, with virtually all existing studies being observational. Inferral from such research is complicated since factors impacting perceived justness may be confounded. For instance, the low-level soldiers who committed the actual crimes and as such have ‘blood on their hands’ may be more remorseful than the distant, higher-ranking orderers. But the low-level soldiers may also be more likely to be tried at local courts. It is therefore hard to assess the impact of one factor net of other factors if they tend to co-occur. Vignette studies overcome such complications. By using a design that is experimental and in which relevant factors are not observed as they occur in the real world, but are manipulated by the researcher, their impact can be validly assessed.

The current study aims to assess the impact of several factors on public perception of appropriateness regarding sentences for perpetrators of international crimes. It investigates how the rank of a perpetrator, presence or absence of apology, and trial location (international tribunal vs. domestic court) influence the public’s perceptions regarding the justness of a certain

---

13 By observational research we mean studies that assess the association between properties of existing cases, properties of perpetrators, other contextual factors and the extent to which sentences are perceived as fair.
sentence imposed on a perpetrator of an atrocity crime. The study is expressly multi-jurisdictional and includes respondents from countries that did not and that did experience (recent) conflict. The experimental design—while not aiming at representativeness—aims to eliminate confounding to assess the impact of trial location, apology and rank of the offender net of respondent sex, age, country of residence and personal conflict experiences.

2 Literature Review

Appropriate punishment for international crimes is generally regarded as promoting—or even a precondition for—peace and stability in post-war societies. If punishments are perceived as appropriate, this could reduce the need for revenge, restore a sense of justice in affected communities, and increase prospects for peace and reconciliation. If punishments are, however, regarded as either too harsh, too lenient, or perceived as biased, the prospects for building a stable and lasting peace may be undermined. Therefore, to encourage peace and reconciliation, it is essential to understand not only what factors influence sentence length in trial procedures, but also the general public’s perceptions and opinions about sentence severity. It is also necessary to determine what factors impact public judgement regarding the appropriateness of sentences.

Research has relatively abundantly investigated the factors that influence sentencing at trials: scholars who study how judges impose sentences in international criminal cases have consistently argued and found that the most important factors that influence sentence length are the gravity of the crime and the culpability of the accused. Meernik examined data on sentences handed down by ICTY judges to understand which factors are the most important in explaining sentences. He found that higher-ranking individuals who exercise power and have command responsibility received longer sentences than lower-ranking individuals. Additionally, judges tended to hand down

16 See Fletcher and Weinstein, supra note 11.
18 Ibid.
20 Ibid., p. 157.
lower sentences to individuals whose cases contained mitigating factors, such as expressions of remorse or a guilty plea. In another study by Meernik and colleagues,21 it was found that punishment is based primarily on the gravity and extensiveness of the crimes as well as the offender’s culpability.22 Similarly, Hola and colleagues examined legal factors influencing sentencing at the ICTY including the number of guilty counts, the rank of the offender, the mode of responsibility, aggravating and mitigating factors (amongst which family circumstances, voluntary surrender and remorse).23 In the model, the number of guilty counts and the rank of the offender were the strongest predictors of sentence length, but some mitigating factors also contributed significantly.24 Furthermore, while investigating judicial decision-making at different international criminal tribunals,25 Meernik found that sentences not only reflect the court’s organisational interests with regards to efficiency, consistency, but also contribution to peacebuilding. Accordingly, plea bargains and expressions of remorse that involved cooperation with the prosecution generally led to lower sentences.26 In the following paragraphs, we will discuss in more detail the three factors that our study focuses on.

2.1 **Rank of the Perpetrator**
Violence committed on such a massive scale that it qualifies as a crime of international concern (i.e., an international crime) generally occurs with the involvement of individuals at the highest echelons of the state or another organisational hierarchy. Without high-ranking individuals who design destructive strategies, implement, and then pursue others to follow them, systematic and widespread violence characteristic for international crimes is unlikely.27 Thus,28 in the case of international crimes, it is commonly accepted that the most responsible are those in high (para)military, police or political positions,29

---

21 Meernik, *supra* note 17.
22 Ibid., p. 70.
24 Ibid., p. 94.
26 Ibid., p. 599.
even though the majority of those on trial may be lower ranking or individuals without any rank.\textsuperscript{30}

In Hola and colleagues’ comprehensive study of ICTY trial sentencing,\textsuperscript{31} the rank of the perpetrator was the strongest predictor of sentence length, net of other case characteristics. This had previously been found in studies conducted by Meernik in 2005\textsuperscript{32} and 2003,\textsuperscript{33} who also found that higher-ranking individuals received longer sentences than lower-ranking individuals. Similarly, Gibson and Gouws\textsuperscript{34} found that South African respondents tended to attribute more blame and responsibility to higher-ranking perpetrators participating in the struggle during apartheid. They were also more likely to advocate for punishment rather than amnesty for higher-ranking perpetrators. Apart from these studies, relatively little research explored whether victims, conflict-affected communities, or the public in general, consider the rank of the perpetrator to be a significant factor in determining appropriate punishment for atrocity crimes.

2.2 Apology

In terms of apologies and remorse, most of the literature related to transitional justice and post-conflict contexts focuses on analysing the importance of official and public apologies. Collective apologies tend to be perceived as a form of symbolic reparation and take the form of formal public announcements admitting responsibility and recognizing that crimes were committed.\textsuperscript{35} Philpot and Hornsey investigated the impact of group apologies on perceived remorsefulness and forgiveness for the offending group in a series of experiments among Australian psychology students. They found that, while the presence of an apology helped promote perceptions that the outgroup was remorseful and the participants were more satisfied with an apology


\textsuperscript{31}Hola et al., supra note 14.

\textsuperscript{32}See Meernik, supra note 17.

\textsuperscript{33}See Meernik, supra note 19.


than without one, the presence of the apology failed to promote forgiveness for the offending group. Apart from such social psychological experiments, the individual dimension of apology in which an individual perpetrator apologises to victims of atrocity crimes has been little studied. Some authors have pointed out that the neglect of remorse and apologies in international criminal justice stems from criminal trials focusing on the assessment of guilt and punishment, rather than on reconciliation processes aimed at healing broken relationships. In addition, criminal proceedings frequently are governed by rigid rules in which there is little room for meaningful apologies. In some cases, victims are not present during trials, there is no direct victim-perpetrator dialogue, and therefore less opportunity for apologies.

Nevertheless, this does not mean that apologies have no relevance and do not occur in international criminal trials. For instance, in studies conducted by Hola and colleagues and Tudor, it was demonstrated that expressions of remorse during trials were relevant factors in sentencing and could eventually be considered as mitigating factors. Hola and van Wijk highlighted that remorse expressions can also be considered when assessing offender rehabilitation during incarceration. It has also been noted that genuine apologies have great potential for healing in contexts where international crimes have been committed. These expressions can signify, not only remorse for a particular crime but also a change of attitude towards the conflict and the victim group. According to Simić and Hola, ‘both remorse and apology can bring catharsis, and reconcile and heal offenders, victims and society’.

When expressed adequately and accepted by victims, they can be an essential

---

38 Ibid.
39 Ibid.
43 Ibid.
44 Simić and Holá, supra note 36, p. 272.
(pre-)condition for social repair.\textsuperscript{45} Despite the healing and reconciling potential of apologies, some challenges may affect the extent to which the public regards apologies as a sufficient reason for sentence mitigation. Remorse and apology as mitigating factors routinely granted in sentencing could create an incentive for offenders to express remorse for opportunistic reasons,\textsuperscript{46} and, at times, expressed remorse is perceived by the public as such.\textsuperscript{47} Consequently, some authors have pointed out that considering all expressed remorse as sincere and as worthy of granting sentence reduction benefits, could potentially harm, not only victims but also the perceived legitimacy of court proceedings.\textsuperscript{48}

However, to date, few studies have empirically analysed the role of apologies in sentencing for international crimes. Research has mainly explored the extent to which judges take these expressions into account when deciding on a sentence and reflected on possible challenges occurring in the process. All in all, not much research has investigated whether victims, conflict-affected communities, or the public in general, consider perpetrator apologies and expressions of remorse relevant when determining punishment for offenders of international crimes.

2.3 Trial Location

Lastly, studies have shown that the type of court, whether national, international or hybrid, is a relevant factor for the opinions and perceptions of communities and victims as to the fairness and justness of a conviction. For instance, Ford developed a (social) psychological model to understand how

\textsuperscript{45} D.F. Orentlicher, \textit{That Someone Guilty Be Published: The Impact of the ICTY in Bosnia}, available online at ictj.org/sites/default/files/ICTJ-FormerYugoslavia-Someone-Guilty-2010-English.pdf (accessed 10 February 2022).


\textsuperscript{48} Simić and Holá, supra note 36. \textit{See} the case of Biljana Plavsic, the former president of Republika Srpska, who was indicted by the ICTY for genocide, crimes against humanity and war crimes. After she pleaded guilty and made a remorseful statement, the prosecutor dropped the genocide charges, and she was sentenced to 11 years in prison. During her time in prison Biljana publicly retracted her statements and stated that she had only expressed remorse to obtain benefits. ICTY, \textit{Prosecutor v. Biljana Plavšić}, Case No. IT-00-39&40/1-S, Trial Chamber, Sentencing Judgment, 27 February 2003.
violence-affected communities perceive international criminal courts. According to the model, if the court’s accusations and sentences contradict a group’s dominant internal narrative about who is responsible for the conflict, the court will be perceived as biased and unfair by that group. This perception allows the members to reject the court’s decisions and preserve their internal narrative. Similarly, drawing on social identity theory, David argued that members of a group welcome decisions of international criminal tribunals as long as they grant them victim status, and reject those decisions that condemn members of the group as perpetrators. Legal decisions taken by non-local courts may therefore trigger protective responses such as denial of harm, or shifting standards of justice. To test the hypothesis, David developed an experimental vignette that manipulated court, i.e., the International Criminal Tribunal for the former Yugoslavia (ICTY) versus Croatian national courts, in-group versus out-group perpetrators, and a five years jail sentence versus no jail sentence. The results of this study indeed demonstrated that the standard of justice is much higher for an out-group perpetrator and significantly lower for an in-group perpetrator and that the conviction of an in-group member is seen as unfair in comparison to the conviction of an out-group member. However, the study did not find any statistically significant evidence in support of the hypothesis which would suggest that international criminal trials are likely to be perceived as more just than national trials. Another study conducted among a sample of victims of the war in Bosnia found that the ICTY was assessed as more just than the Court of Bosnia and Herzegovina. This implies that the type of court (i.e., domestic vs. international) adjudicating the case may be a relevant factor in the public’s evaluation of the justice process.

Saxon assessed how three different national communities in the former Yugoslavia, the Serbs, the Croats and the Bosniaks/Muslims, perceived the ICTY—either as a success or a failure with regards to achieving peace and justice. According to this study, perceptions of the ICTY are diverse, varying from

50 Ibid., p. 410.
51 David, supra note 10, p. 481.
52 Ibid., p. 489.
53 Ibid., p. 492.
55 Ibid.
community to community, but also within communities. For example, Serb respondents opposed the work of the ICTY, while within the Bosniak/Muslim and Croat communities opinions varied: for some, the ICTY represented hope, while for others the ICTY’s decisions were of little to no relevance. Similarly, Arzt examined local perceptions of two international tribunals, the ICTY and the Special Court for Sierra Leone, from three perspectives: legitimacy, impartiality, and the effect of the courts on the reconciliation process. The opinions were mixed. While the perceptions of the ICTY among Bosnian Muslims were much more positive than among Serbs or Croats in the region, in Sierra Leone the perception of the Special Court depended more on details and implementation than generic legitimacy.

All in all, the limited literature has not given definitive answers on the perceived legitimacy of sentences handed down by international versus national courts. To fill this gap in the literature, the current study aims to address the perception of justness in international trials in comparison to national trials.

3 Research Questions

The aim of our study is to assess the impact of perpetrator rank, perpetrator apology and trial location on the perceived fairness of sentence length. To this end, our study will focus on the impact within a given vignette of perpetrator role (commander versus soldier, where we expect respondents to desire a longer sentence for the commander), an apology by the perpetrator (where we expect an apology to promote approval of shorter sentences) and trial location (where we do not have an expected direction of the effect), all randomly varied.

We first asked respondents to what extent they assessed, following David, as ‘just’ a sentence length of eight years given in a vignette (which constitutes a realistic, average sentence for the given vignette case that is based on a real case adjudicated following the war in the former Yugoslavia), and this variable we will refer to as ‘justness’. Next, we asked them to explain their motivation

57 Ibid., p. 564.
59 Ibid., p. 232.
60 Ibid., p. 233.
61 David, supra note 10.
62 In the literature different terms for similar outcome measures are encountered, such as ‘appropriateness’, ‘fairness’, ‘justness’, or even ‘justice’—although that generally has a more
for such a judgement. Lastly, we asked them what they would regard as the appropriate sentence length. It is this last measure that is our dependent variable of interest, for which we will investigate the impact of apology, trial location and rank of the perpetrator on sentence length.

4 Method

4.1 Vignette

A vignette describing a case of a war crime was designed in order to investigate the way in which the manipulated factors of the case (rank, apology and trial location) influence desired sentence length. The vignette described a case of a paramilitary group that attacked a village and took civilians to locations where they were illegally detained in poor conditions. During their stay in the detention facility, some civilians were mistreated which resulted in two victims ultimately dying of their injuries. The vignette then described that one of the responsible perpetrators was accused of illegal detention, torture and death of civilians and tried in a court of law. The perpetrator was found guilty and given a prison sentence of eight years.63 The described case was based on real events (during the Bosnian war) while keeping the descriptions of maltreatment more abstract and the location of the crime unspecified. This was

legal connotation. In preliminary analyses the face validity of the term ‘justness’ emerged as most appropriate as it encapsulates both notions of ‘justice’, and ‘appropriateness’ or ‘fairness’.

63 During the recent civil war in country ‘A’, a number of paramilitary groups were active. These groups consisted of men mostly, although a number of women took part as well. The groups were organized hierarchically, with senior ‘commanders’ in charge and regular, lower ranking ‘soldiers’. In the evening hours of 1 May 2012, members of such a group carried out an armed attack on a civilian population in a small village. Around 1000 civilians were attacked: many were able to flee, but around 300 civilians, including elderly and children, were taken from their homes, and illegally detained. Conditions were bad in the detention facilities. David, a commander in this rebel group, acted against some civilians on a daily basis where they were beaten and kicked. Some men and women were tortured; in total two victims died of their wounds. In mid-May the illegal detention was discovered by the International Committee of the Red Cross and the situation ended. David stands trial at the State Court of in the capital of country A for the illegal detention, torture and death of civilians. David is currently working as an administrator in a university hospital, he is married and has two children. During the trial, David apologized to the surviving victims and family members of the deceased victims. David was found guilty for personally committing torture and illegally detaining civilians, and for aiding and abetting in the two deaths. He was sentenced to 8 years imprisonment.
done to prevent potential secondary victimisation and to minimise the possibility of biased answers or sensitivities in some of the participants, especially those who might have previously experienced conflict situations.

Multiple versions of the vignette were created across which we randomly varied, amongst others, the rank of the perpetrator, the trial location, and whether or not the perpetrator apologised for their actions. The vignettes were originally created in English but were translated (and back-translated) to languages native to the researchers’ respondents (e.g., Croatian, Spanish, Dutch). Before administering the vignettes to the respondents, they were piloted in the Netherlands and several other countries. The process showed that the vignettes did not cause any ambiguities or sensitivities.

The description of the case was followed by questions for the respondents. Firstly, respondents were asked to express the extent to which they perceived the court-assigned sentence as just and to offer a qualitative explanation for their answer. If they perceived the sentence length as unjust, they were asked to state the length of the sentence (in months) which they regarded as more suitable for the case at hand; if they did not believe it needed to be changed, the respondents’ desired sentence length was coded as conforming the case description. In this study, we will analyse the latter two variables: preferred sentence length and people’s responses to the sentence handed down in the hypothetical vignette. After this, respondents were inquired about their sex, age, and any personal experience with conflicts such as (civil) wars.

4.2 Procedure
The 570 responses were collected by 57 fieldworkers in multiple countries. The sample constitutes a convenience sample as each of the fieldworkers gathered ten respondents willing to participate in the study. Fieldworkers were asked to ensure a reasonable spread over age, gender, and occupational class across their samples. Because the study was realised in multiple countries that had varying COVID-19 measures in place, flexible methods of administration were employed by the fieldworkers. The responses were gathered through video-call, telephone conversation, self-administration by respondents via email or in person at one and a half metres distance. To reduce the risk of confounding, each fieldworker was tasked with administering two randomly allocated different vignettes.

Before proceeding to the case description, study participants were presented with introductory information on the goals and procedure of the study, informed about voluntary participation as well as anonymity and confidentiality of their responses. It was also emphasised to them that the case, while
based on true events, was hypothetical. After the introductory note, they proceeded to read the description of the case and answer the posed questions.

4.3 Sample
A total of 570 respondents agreed to participate in the research. Respondents hailed from a large number of different countries, 39 in total. The largest groups were respondents from the Netherlands (45%), Germany (7.7%), Croatia (5.4%), Italy (5.4%), France (5.3%), Israel (1.9%) and Colombia (1.5%); for 46 respondents country information was missing. Most data were collected through questionnaires filled out during face-to-face interviews (44.2%); 22% of interviews were conducted by telephone, 27.7% via Zoom or some other type of video-call while only 5% of the questionnaires were administered through email and filled out personally by respondents.

4.4 Analysis Method
We first describe our findings using simple graphs. Next, we analyse the quantitative impact of the manipulated variables on sentence length using a factorial Analysis of Variance (ANOVA), with analyses carried out using SPSS version 24. Subsequently, we analyse the qualitative remarks respondents gave as a rationale for their choice of sentence length. For this we coded the qualitative answers by using Atlas-ti and a mixture of grounded and theoretical coding: we paid specific attention to any remarks made regarding the trial location, the rank of the perpetrator and apology. Additional remarks were coded using an open and adaptive coding frame. After a first round of coding, a second researcher checked and revised the coding.

5 Results

5.1 Descriptives
Valid responses on gender were collected from a total of 563 respondents: 54.2% were women, 45.8% were men. Respondents’ average age was 37.9 years, with the youngest respondent being 18 years and the oldest 94 years of age (Standard Deviation: 17.7). Out of 566 respondents who provided information about their experience with conflict, 10.2% stated they had personally experienced a conflict such as a (civil) war.

We first describe frequencies (on the y-axis) of given sentence lengths (on the x-axis) as per the variables rank perpetrator (Figure 1), apology (Figure 2) and trial location (Figure 3). Quite a number of respondents gave as a response for desired sentence length ‘life sentence’, which we recoded to 65 years.
Next, we investigate to what extent the rank of the perpetrator, apology, and trial location influence sentence length. In doing so, we were faced with two complications. First, even though field workers administered a random sequence of vignettes, we may not consider our design as completely balanced. We, therefore, analyse the impact of the various factors in a multivariate model that includes, apart from the independent variables, a number of respondent characteristics to control as much as possible for any confounding. Second, 18% of respondents did not hand down a determinate sentence, but stated that the perpetrator should receive a ‘life sentence’. As this is not a determinate sentence, we ran our analyses three times: once where we had coded ‘life’ as 65 years (which would mean that the perpetrator would be incarcerated until a minimum age of 91, and expresses that life is a heftier sentence that a maximum determinate sentence), once where we coded life to 50 years which was the maximum determinate sentence handed down, and last with a ‘symbolic’ value of 100 years. The results from all analyses were substantively the same. We report the findings for the first option.

Table 1 gives the results from the ANOVA which showed that, of the manipulated (fixed) variables, only apology had a significant impact on sentence length. Figure 1 shows that soldiers in the vignette tend to less often receive life sentences and slightly lower determinate sentences than commanders. As Figure 1 shows (and the following figures too), the variable sentence length is heavily skewed, also without life sentences coded as ‘65 years’.

5.2 Quantitative Analysis of Impact of Manipulated and Respondent Factors

Figure 1 shows that soldiers in the vignette tend to less often receive life sentences and slightly lower determinate sentences than commanders. As Figure 1 shows (and the following figures too), the variable sentence length is heavily skewed, also without life sentences coded as ‘65 years’.
length: apology reduced sentence length by approximately three point four years. While the impact of the rank of the perpetrator was in the expected direction—with soldiers receiving shorter sentences than commanders—the difference was not significant. Sentences handed down in the vignettes where the trial had taken place at the ICC were generally shorter. Of the additional (random) variables (age recoded into six equally large age groups, sex, country recoded into 13 main geographical regions, and experienced conflict) included in the ANOVA, the region where the respondent hailed from had a very strong impact. Respondents from regions or countries that had recently experienced conflict, such as Croatia, handed down long sentences. One reason for this may be that these respondents had themselves experienced conflict; however, the variable that indicates whether this was the case, was not significant. Also, when the variable measuring respondent region was removed from the analysis the variable that indicated whether respondents themselves had experienced conflict did not become significant. The impact of the region is therefore not due to confounding with conflict experience.

5.3 Qualitative Analysis of Rationales
Next, we analysed the qualitative responses that had been given as respondents’ reasons for a certain given sentence length, either because they were fine with the eight years in the vignette or because they thought it should be longer.
which was mostly the case) or shorter. A total of 238 times we found that respondents had referred in their answers to one or more of the manipulated factors (apology, rank perpetrator and trial location).

In these answers, the apology was by far the most prominent (it had been mentioned 76 times). Only three respondents raised doubts or stated that it should be clear that the apology should be genuine. In the remaining answers,

**FIGURE 3** shows a somewhat mixed picture: on the one hand, perpetrators who are sentenced at the ICC appear to receive fewer life sentences, but their determinate sentences appear somewhat longer than the sentences of those who are sentenced within the country where the atrocity crime was committed.

**TABLE 1** Factorial ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>ss</th>
<th>df</th>
<th>ms</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country trial</td>
<td>279.56</td>
<td>1</td>
<td>279.56</td>
<td>0.692</td>
<td>0.406</td>
</tr>
<tr>
<td>Apology</td>
<td>2635.86</td>
<td>1</td>
<td>2635.86</td>
<td>6.522</td>
<td>0.011</td>
</tr>
<tr>
<td>Rank</td>
<td>314.87</td>
<td>1</td>
<td>314.87</td>
<td>0.779</td>
<td>0.378</td>
</tr>
<tr>
<td>Region</td>
<td>28454.85</td>
<td>12</td>
<td>2371.24</td>
<td>5.876</td>
<td>0.000</td>
</tr>
<tr>
<td>Age</td>
<td>1940.82</td>
<td>5</td>
<td>388.16</td>
<td>1.268</td>
<td>0.442</td>
</tr>
<tr>
<td>Sex</td>
<td>512.64</td>
<td>1</td>
<td>512.64</td>
<td>1.268</td>
<td>0.261</td>
</tr>
<tr>
<td>Conflict</td>
<td>11.61</td>
<td>1</td>
<td>11.61</td>
<td>0.029</td>
<td>0.865</td>
</tr>
</tbody>
</table>
54 respondents mostly indicated that if no apology had been given, this should imply a longer sentence. We quote a few answers: ‘Lack of remorse deserves longer sentence’. ‘Because [the perpetrator] does not show remorse and the crime is severe, there is a need for an appropriate punishment’. ‘[The sentence is] long, but since not apologised it’s suitable’. ‘[The perpetrator’s] lack of empathy and apology should also add time’.

On the other hand, the apology by the perpetrator much less often made respondents propose a shorter sentence or deem the sentence too long (13 times). Again, we quote a few answers: ‘[The perpetrator] showed remorse, asked for forgiveness and recognized [the suffering of the victims]’. ‘Too short for two deaths, although the apology helps’. ‘[It is] already ten years later, [the perpetrator] offered his excuses and has also built a new life’.

It appears, therefore, that apology was not considered a mitigating factor in the situation in the vignette, but that rather the absence of an apology was considered an aggravating factor. We see a somewhat similar pattern emerge if we look at the qualitative reasons given that referred to the rank of the perpetrator. Here we see that a high rank was considered a reason for a longer sentence (33 times) than given in the vignette, but a low rank was less often considered a mitigating factor (eight times). Again, we give several quotes: ‘As a commander [the perpetrator] carries more responsibility for the acts’. ‘Since [the perpetrator] was commander of the group they are responsible for the full group’. ‘[The perpetrator] was the commanding officer, most likely very well educated and should be held accountable for [the crimes]’.

Quite a few respondents even stressed that as commander, the perpetrator was responsible for the actions of the entire group. On the other hand, for the soldiers in the vignette, respondents regularly stressed that the perpetrator was probably under orders to commit the crimes: ‘As a simple soldier [the perpetrator] probably received the order for the actions’. ‘[The perpetrator] was a soldier following orders and the leaders are more war criminals’. ‘[The perpetrator] wasn’t in command’.

Trial location was rarely referred to (two times) and we will not elucidate these answers further. All in all, we see that the qualitative analysis clearly underlines our quantitative finding that apology is a prominent factor impacting respondents’ views on what is a just sentence for the described crimes. The qualitative data nuance the quantitative findings, however, in that we see that the impact of apology is not balanced: it is mainly the absence of an apology that respondents find warranting a longer sentence.
6 Discussion and Conclusion

In order to investigate factors impacting the perceived justness of sentences assigned to perpetrators of international crimes, we employed a mix of quantitative and qualitative analysis methods. Opting for an experimental vignette study enabled us to quantitatively assess the impact of perpetrator rank, perpetrator apology and trial location on preferred sentence length, net of confounding factors that generally make causal inference harder when studying sentencing in empirical reality. Additionally, by attaching open-ended questions to the vignette, we were able to extract and analyse motivations and reasons behind people’s perceptions of what would constitute an appropriate sentence in this case.

We found that only a perpetrator’s apology significantly affected sentence length. While commanders were more likely to receive longer sentences than soldiers and sentences handed down for cases processed at the international court were generally shorter, these differences were not significant. This finding is especially interesting as it appears to paint a different picture from the evidence found so far in observational studies on judicial reasoning. That is, while those studies examining judicial decision-making consistently found that perpetrator rank significantly impacted sentence length, this was not upheld for the public’s assessments. This hints at potential differences in how ‘doing justice’ and ensuring accountability for atrocious international crimes is perceived by judges and experts on the one hand, and the affected population or the general public on the other. Future studies should explore these differences more in-depth in order to find a way to understand the gap between them. Understanding this gap (and ultimately filling it) is essential as, to be successful in their aim to achieve accountability and contribute to peace and reconciliation, it is important that the judicial proceedings are seen as just, not only by judges and lawyers but also by the affected community or the general public.

Additionally, we found that the region the respondents hailed from had a very strong impact on their perceptions of what constituted an appropriate sentence length. Interestingly, our explanation that respondents coming from (recent) post-conflict countries (e.g., Croatia or Colombia) would tend to give longer sentences because they experienced the conflict, was not confirmed. Thus, this country-effect is possibly an effect that is rather attributable to general punitive norms. This should also be explored in more detail in future studies to understand in what way and why it so strongly impacts people’s perceptions regarding punishment for international crimes.
Respondents gave a variety of arguments when explaining their decisions regarding the appropriate sentence length for the crime described in the vignette. Qualitative analysis revealed that the perpetrator’s apology was mentioned most often. The results clearly implied that the apology was not so much considered a mitigating factor when deciding on an appropriate sentence, but rather that the absence of apology was considered an aggravating factor. These results are largely in line with our quantitative results. They add on, in an interesting way, to the existing literature on the relationship between apologies and sentencing. That is, while the existing literature shows that apologies and remorse can serve as mitigating factors in sentencing, our study has shown that it may be the absence of it that is considered particularly heinous and actually an aggravating factor.

Overall, trial location did not emerge as a relevant factor in the quantitative or qualitative analysis. This might be explained through either or all of the following arguments. The fact that trial location did not emerge as a significant factor might simply mean that it is not relevant to the respondents whether the case is tried at the international or a domestic court as long as the perpetrator is held accountable. It could also be possible that our respondents did not have enough knowledge about the (inter)national criminal tribunals and (inter)national criminal justice systems to be able to make relevant inferences when presented with information regarding trial location. In order to explore this further, future research could consider exploring perceptions of people with knowledge or personal experience (e.g., coming from a country whose citizens were on trial for international crimes) regarding trials for international crimes. Alternatively, future (vignette) studies could put more emphasis on assessing the impact of trial location with regards to international crimes by giving more contextual or theoretical information about (inter)national criminal justice systems.

While the inferences derived from this empirical study provide the field of transitional justice and empirical scholarship on international criminal law with relevant information, some methodological limitations should be addressed. To begin with, during the data collection, the researchers noticed that some respondents failed to notice or pay attention to factors that we manipulated. While we attempted to make participation attractive by keeping the vignette short, perhaps we should have made the factors we were manipulating stand out more. For instance, in the vignette without an apology by the perpetrator, we did not highlight this by including a sentence like ‘the perpetrator did not apologise’, but simply omitted the sentence mentioning

64 See Hola, supra note 39. See also Tudor, supra note 40.
the apology. Similarly, when manipulating the perpetrator rank, we simply replaced ‘soldier’ with ‘commander’, without any emphasis on this role. It should be remarked that, even so, ‘apology’ did emerge as a predictor of sentence length, and a prominent theme in the qualitative analysis. Second, as the sample of respondents was not representative of the general population in each of the countries where they hailed from but a convenience sample, generalizability of the findings (although we attempted to ensure a good spread of occupational class, age, and gender) may not be assumed. Third, the administration of vignettes, while randomly assigned, did not emerge as completely balanced in the practical realisation of data collection. While we statistically controlled for this as much as possible by analysing the impact of the various factors in a multivariate model that includes possible confounders, inference should be careful. Fourth, it is possible that the offences described in our vignette were so severe that this may have prevented our manipulated factors (the independent variables) from affecting respondents’ perceptions and opinions. That is, it may be that the respondents were merely focused on ensuring retaliation and punishment of the perpetrator—regardless of any other factors we mentioned. If we had composed the vignette around a less severe crime, a factor like the perpetrator’s rank would potentially have a more decisive impact on sentence length.

Finally, it has to be emphasised that judicial proceedings against perpetrators of international crimes are complex, long-lasting, and politically charged, and that the ramifications of such crimes are severe and pervasive. Accordingly, those events can hardly be summarised to less than a page of text, and any consideration of what is just or appropriate can hardly be determined in a few minutes of participation in a study. In short, our study, even though it scores high on internal validity, might also be said to have low ecological validity. People’s perceptions about such complex events in real life will be influenced by many other factors (e.g., official war narratives, media coverage, revisionist narratives, etc). Nevertheless, the inferences deduced from this and similar studies can be significant and valuable for the assessment of past and development of future transitional justice efforts.

Acknowledgements

We thank two anonymous reviewers and Barbora Hola for their valuable comments on previous versions of this text. The study was designed and executed as part of the course Research Methods for International Crimes that is a part of the Master’s programme International Crimes, Conflict, and Criminology.
at Vrije Universiteit, Amsterdam. All students in the course contributed to the study materials, piloting of the design, data collection, along with analysis and interpretation. Margareta Blažević and Diana Bociga Gelvez were responsible for the literature review and the general discussion, Catrien Bijleveld conducted the quantitative data analysis, Diana Bociga Gelvez conducted the qualitative data analysis; all three wrote the text while Mirza Buljubasic contributed to the vignette design and the editing process. All other contributors are in alphabetical order: Alexandrah Bekker, Margareta Ana Baksa, Anna Ballauf, Marielotte van Ballegooijen, Tim Bekkers, Frans van der Bloom, Milena Borger, Ella Boulam, Annelize Bourcier, Vera Breukink, Lidewij de Bruijn, Lena Chouchan, Bindi Corpe, Emerson Dahl, Francisca De Castro Forero, Mariachiara Dellora, Sophie van Denderen, Stijn Dormans, Ebba Fernholm, Michal Fisher, Inne Flies, Billie Goode, Fabio Guadalupi, Ruben van Hamersveld, Margot van Heesbeen, Tessa Huizing, Ruth Juursema, Tessa Kooi, Nena Kos, Romee Kruyt, Lieke Kuiper, Sara Kulić, Sofie Lemeris, Aimée Laure, Fernanda Luft, Miloš Maggiore, Beatriz Eugenia Mayans Hermida, Sarah van de Meent, Esther Millet, Ayla Molenkamp, Bianca Montesanti, Verena Muckermann, Coralie Niggeler, Madison Pelletier, Nini Pieters, Nathalie Prange, Petra Regeni, Rowie Rensink, Tess Rijk, Marit Rutten, Beatrice Schutzbach, Pia Sienz, Lily Steffen, Linde Trommel, Angelique Truijens, Tess Veldhoven, Daniël Versteeg.