

'VOX POPULI, VOX DEI'?

A COMPARATIVE INVESTIGATION INTO THE (UN)FAIRNESS OF
THE JURY TRIAL IN THE BRITISH AND ITALIAN LEGAL SYSTEMS

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Cristina d'Aniello

Department of Criminology

University of Leicester

2018

Abstract

'*Vox Populi, Vox Dei*'? A comparative investigation into the (un)fairness of the jury trial in the British and Italian legal systems.

Cristina d'Aniello

Behind the closed doors of the deliberation room, jurors determine another person's fate. It is believed that, to grant fairness to trials, the decision of some judicial cases has to be referred to a panel of impartial peers. Yet, the occurrence of miscarriages of justice involving incorrect jury verdicts demonstrates that the jury system may be failing to respond to those democratic needs that constituted the foundations of its introduction. In an effort to identify the causes of the malfunctioning and to propose solutions, this thesis has investigated the matter through a comparative approach that looks at two crucial differences between British and Italian juries: the presence/absence of professionals (judges) on jury panels and the presence/absence of a requirement for juries to justify their verdicts. Far from being mere procedural aspects, these characteristics play a crucial role in the deliberation process, as this research found through an analysis of the results yielded by two interconnected empirical studies: interviews with Italian judges and mock jury experiments. Results from the studies suggest that both jury composition and motivated verdicts have an impact on juries' behaviour, errors, and deliberation dynamics. Beneficial and detrimental effects of the two variables were considered in order to suggest solutions for an improvement to the functioning of jury trials. Accordingly, the aid of a professional juror, purposely trained to instruct and direct (not influencing) the panel of peers, could improve legal fairness of deliberations. Additionally, motivated verdicts should be required, since they increase jurors' tendency to provide legally-oriented decisions. Given the high real-world impact of the matter, the implementation of these and further research suggestions is crucial to move towards the 'fair trial' that the jury system promised to grant.

Acknowledgements

There are many people without whom this research project could not have been completed and to whom I am sincerely, deeply grateful.

To Prof. Lisa Smith, my first supervisor (and favourite superhero!), for everything she taught me. If I ever met someone who made me believe that (almost) everything is possible, that was her. The joy and enthusiasm she has for research have been contagious and motivational for me, even during the toughest times in my PhD journey, which she helped me face with her constant encouragement. Thank you, Lisa, for supporting me and this research project with great energy and endless patience. Having you as a supervisor and as a mentor has been a never-ending source of inspiration and drive.

To Dr. Clare Gunby, for her huge, priceless contribution to the supervision of this research. Her meticulous comments and incredibly valuable suggestions have constantly stimulated further reflections and generated thought-provoking conversations, which I deeply enjoyed and will bring with me far beyond the completion of this piece of research. Thank you, Clare.

To the Italian judges, and the mock judge and jurors, who participated in the studies conducted for this research: without your contribution, this project would never have been completed. Thank you all very much!

To my family, as always, for everything. For supporting me whenever I need it, and asking for my support whenever they need it. For always being proud of me, and always making me proud of them. For being with me wherever I go, and taking me with them wherever they go. For I know that our love is and will always be my greatest success in life. Dad, Mum, Rosy, Checco, I love you all.

To Massimiliano, for always being... there.

To Anne, for teaching me all that really matters in life, like children often do.

And to Davide, for holding my hand every step of the way, ever since we met, and I'm sure, somehow, even before. Always.

Table of Contents

Abstract.....	1
Acknowledgements	2
Table of Contents.....	3
List of Tables and Figures.....	10
INTRODUCTION.....	11
CHAPTER 1: LITERATURE REVIEW	15
1.1 Introduction	15
1.2 The Jury Trial: exploring the past to understand the present	16
1.2.1 The origins of the jury: the introduction of the jury in ancient Greece and Rome.....	16
1.2.2 The origins of the jury: the introduction of the jury in England.....	18
1.2.3 The origins of the jury: the introduction of the jury in Italy.....	19
1.3 World jury systems’ differences: not only formal distinctions.....	21
1.4 Are some jury systems “better” than others?	24
1.5 Jury decision making – a complex and unique task.....	26
1.5.1 Decision-making.....	27
1.5.2 Decision-making under ignorance and/or uncertainty	29
1.5.3 Individual (jurors) decision-making – theoretical and empirical perspectives	32
1.5.3.1 Mathematical/Probabilistic Approaches.....	33
1.5.3.2 Cognitive Approaches.....	34
1.5.3.3 Jurors’ Errors.....	39

1.5.3.4	Further reflections	42
1.5.4	Group (jury) decision-making – theoretical and empirical perspectives	43
1.5.4.1	Leadership	45
1.5.4.2	Conformity	48
1.5.5	Experts’ opinions influence	51
1.5.6	Explanation-based reasoning	53
1.6	Conclusion	55
CHAPTER 2: METHODOLOGICAL CONSIDERATIONS		56
2.1	Introduction	56
2.2	The overall methodological approach: mixed methods design	57
2.3	Research questions	58
2.4	Research design	61
2.4.1	Study 1: Italian professional judges interviews	61
2.4.2	Study 2: Mock juries experiments	65
2.5	Concurrent and Sequential mixed methods designs	72
2.6	Conclusion	73
CHAPTER 3: ITALIAN PROFESSIONAL JUDGES INTERVIEWS: A look inside the deliberation room		75
3.1	Introduction	75
3.2	Method	77
3.3	Findings and Discussion	83
3.3.1	Judges’ role on the jury panel	84

3.3.1.1	Working experience on jury panels.....	86
3.3.1.2	Training/Professional Development.....	86
3.3.1.3	Judges' personal interaction with the jurors.....	87
3.3.1.4	Judges' behaviour/actions prior to the beginning of the trial	90
3.3.1.5	Judges' behaviour/actions at the outset of the deliberation.....	93
3.3.1.6	Judges' behaviour/actions during the deliberation (how they manage/direct the discussion)	95
3.3.1.7	Judges' behaviour/actions when jurors do not remember relevant facts or evidence.....	97
3.3.1.8	Judges' behaviour/actions when jurors misinterpret forensic evidence	99
3.3.2	Motivation	102
3.3.2.1	Judges' attitude towards the need to motivate verdicts	103
3.3.2.2	How do judges manage to write a motivation for a verdict on which they do not agree?.....	105
3.3.2.3	When is the motivation (reason for choices) asked?	107
3.3.3	Judges' perception of their role.....	109
3.3.3.1	Influence	110
3.4	Conclusion.....	114
CHAPTER 4: MOCK JURY EXPERIMENTS: Comparing British and Italian (mock) juries		116
4.1	Introduction	116
4.2	Method	117
4.2.1	Participants	117
4.2.1.1	Nature of the sample and recruitment – Lay participants	117
4.2.1.2	Nature of the sample and recruitment – Professional participant	118

4.2.1.3	Sample Size	120
4.2.2	Design.....	121
4.2.3	Materials	122
4.2.4	Procedure.....	124
4.2.4.1	Pilot and Mock Juries in the British condition.....	125
4.2.4.2	Mock Juries in the Italian condition	127
4.2.5	Ethics.....	128
4.3	Analysis	129
4.3.1	Stage 1: Identification of themes and coding process.....	129
4.3.2	Stage 2: Questionnaires' responses – preliminary exploration.....	131
4.3.3	Stage 3: Video deliberations – first view	131
4.3.4	Stage 4: Video deliberations – second view	132
4.3.5	Stage 5: Video deliberations – third view: double-check with questionnaires	135
4.3.6	Stage 6: Video deliberations – final view and collection of data from questionnaires	135
4.3.7	Qualitative Analysis.....	136
4.3.8	Quantitative Analysis	137
4.4	Findings and Discussion	137
4.4.1	Mock Juries' Deliberations – Qualitative Content Analysis.....	138
4.4.1.1	Jurors' Behaviour.....	139
4.4.1.1.1	Approach to deliberation.....	139

4.4.1.1.2	Narrative Construction.....	142
4.4.1.1.3	Confirmation Bias.....	147
4.4.1.1.4	Leadership.....	150
4.4.1.2	Errors	154
4.4.1.2.1	Memory Errors.....	154
4.4.1.2.2	Evidence-related Errors	161
I.	Misunderstanding/Misconception	164
II.	Overestimation.....	166
4.4.1.2.3	Law-related Errors.....	169
I.	Misunderstanding of Reasonable Doubt.....	170
4.4.1.3	Differences between the two conditions.....	174
4.4.1.3.1	Presence of the mock judge.....	175
I.	Advantages	175
–	Summary report	175
–	Questions.....	176
–	Clarifications	177
–	Accuracy.....	178
II.	Disadvantages.....	179
–	Frequent interruptions	180
–	Influencing remarks.....	181
–	(Too) Early opinion disclosure	181
–	Control over topics addressed.....	182
4.4.2	Mock Juries’ Deliberations – Quantitative/Statistical Analysis	183
4.4.2.1	Sample	183
	Gender, Age and Nationality.....	183

Occupation and Jury Service	184
British condition and Italian condition (composition)	185
4.4.2.2 Deliberation duration	185
4.4.2.3 Verdict preferences	186
4.4.2.4 Conformity	188
4.4.2.5 Confidence	192
4.4.2.6 Perception of verdict fairness	193
4.4.2.7 Leadership Perception	195
4.4.2.8 Need for/usefulness of a judge or someone to direct the discussion	198
4.4.2.9 Motivation	200
4.5 Conclusion	203
CHAPTER 5: FINAL DISCUSSION	204
5.1 Introduction	204
5.2 Impact of composition and motivated verdicts on jurors' behaviour	205
5.3 Impact of composition and motivated verdicts on jurors' errors	209
5.4 Impact of composition and motivated verdicts on jury deliberation dynamics	211
5.5 Lay vs Mixed Juries – is there a “winner”?	215
5.6 Guilty or not guilty – what if we asked ‘why’?	217
5.7 Contributions, strengths, limitations and future research suggestions	219
CONCLUSION	228
APPENDIX A: Italian judges' interviews – Informed consent form	231
APPENDIX B: Italian judges interview questions	234
APPENDIX C: Mock jurors – Informed consent form	236

APPENDIX D: Mock jury experiments flyer	239
APPENDIX E: Crime case scenario	240
APPENDIX F: Crime case presentation	241
APPENDIX G: Prosecution and Defence closing arguments	242
APPENDIX H: Demographic questionnaire and notes	246
APPENDIX I: Pre-/Post-deliberation questionnaires	248
BIBLIOGRAPHY	253

List of Tables and Figures

Table 3.1: Italian Judges' Interviews – Topics and Themes.....	84
Table 4.1: Qualitative Content Analysis – Coding Template.....	138
Table 4.2: Confidence pre-/post-deliberation across conditions.....	193
Table 4.3: Observed vs. Perceived Leadership in British condition.....	196
Figure 2.1: Study 1 – Research Questions.....	59
Figure 2.2: Study 2 – Research Questions.....	60
Figure 4.1: Deliberation Duration.....	185
Figure 4.2: Pre-deliberation Verdict Preferences.....	188
Figure 4.3: Rape Verdict Conformity.....	190
Figure 4.4: Murder Verdict Conformity.....	190
Figure 4.5: Rape Verdict Conformity by condition.....	190
Figure 4.6: Murder Verdict Conformity by condition.....	190
Figure 4.7: Need for/Usefulness of Judge.....	199
Figure 4.8: Need for/Usefulness of someone to direct.....	199
Figure 4.9: Type of Motivation.....	201

INTRODUCTION

‘How would you like to have your fate decided by twelve people who weren't smart enough to get out of jury duty?’ (Shuman and Champagne, 1997, p.249). Jurors, the ultimate expression of the popular sovereignty in the administrations of justice, are in reality severely criticised for the malfunctioning of the system within which they work. The jury system, introduced as a democratic guarantee (Hans, 2008), has in fact often failed to grant “fair” trials and jurors have frequently taken the blame for this failure. At closer inspection, it seems that to blame the jurors themselves is somewhat like blaming the messenger for delivering bad news (Kassin and Wrightsman, 1988), especially when considering that the task they are required to perform is, in fact, not an easy one (Arkes and Mellers 2002). These people, without receiving any form of training in the discipline of law or information about the logic that stands behind legal matters, have to decide criminal cases on the basis of the evidence presented at trial, which in fact entails making evaluations and consequent choices on the basis of very complicated and confusing facts (Vidmar, 2000).

Scholars have long shown interest in the jury task, because, unlike some other comparable decision-making processes, jury deliberations have serious consequences in legal matters (Hastie, 1994). Accordingly, the need to conduct research on the functioning of the jury trial is strongly felt. Therefore, research has focused on how jurors make their decisions, from both a theoretical and an empirical perspective. However, given the prohibitions set to grant secrecy to the jury deliberation and avoid pressure on its actors, only a small proportion of research on the jury system has achieved the result of investigating the attitudes and behaviour of actual jurors in situ (Tinsley, 2001).

This thesis conducted an investigation into the topic, which intended to examine unexplored aspects of it, whilst attempting to overcome limitations that have generally affected previous empirical endeavours. A consideration of the (mal)functioning of the trial by jury in two different criminal justice systems – British and Italian – has led to the realization that new insights would be gained through a comparison of elements of

difference, which, despite being procedural, may well affect the jury deliberation task. In particular, the Italian jury system established the duty for juries to provide a motivation for their decision as an essential and compulsory requisite for the validity of the verdict (Apa and Cantarini, 2011), and also requires the presence of professionals (judges) on the jury panel (two out of eight members). The absence of corresponding rules in the British system has allowed the comparison that this thesis presents.

Two interconnected studies were designed so that each study's results, besides having their own value, would simultaneously become methodological elements to use in the other study. In Study 1 (Chapter 3), interviews with Italian judges who work on juries were conducted in order to better understand the role of these professionals. The results were then used in Study 2 (Chapter 4), in which mock jury experiments were conducted in order to compare elements of the British and Italian systems. The ultimate scope of the overall analysis was to better understand whether the need to provide motivated verdicts and the presence of judges on jury panels exert an impact on deliberations, in the belief that, by exploring this uncharted territory, novel results on the topic could contribute to suggest solutions aimed at the achievement of a fair(er) jury trial.

This thesis is composed of five chapters, which are structured as follows:

Chapter 1 provides a brief historical overview of the introduction of the jury trial, aimed at highlighting how the fundamental reasons that are at the basis of the institution governed it throughout its evolution. An overview of the modern jury trial functioning is provided with specific emphasis on various compositions (purely lay or mixed) and whether juries need to provide motivated or unmotivated verdicts. Following, a review of the literature on decision-making is provided, with emphasis placed on specific topics of relevance to jury decision making. Through a review of broad aspects of decision-making and specific aspects of jury decision-making in both its individual and collective dimension, the chapter concludes with reference to the

two variables under examination in this thesis: presence/absence of judges on the panel and presence/absence of the requirement for motivation.

Chapter 2 provides general methodological considerations for the studies conducted in this thesis, with the main aim to provide justifications for methodological choices and highlight the interconnected nature of the two studies. In this context, reference to the overarching research question and sub-questions is made, both studies' designs are outlined, and benefits and limitations of the employment of mixed methods are highlighted.

Chapter 3 presents the first study in which interviews with Italian judges who work on juries were conducted and thematically analysed. Further specific reflections on methodological choices, along with benefits and limitations of the chosen methodological tool were provided to complement the general observations made in Chapter 2. Throughout the chapter, themes which emerged during the interviews are addressed and considerations regarding the role of the interviewed professionals are made, with specific attention to the implications that their behaviour may have on deliberations. A certain degree of mismatch between the actual role of the judges and their perception of it is discussed in the context of some contradicting remarks. Finally, reflections and conclusions are proposed regarding judges' attitudes towards the motivation for verdicts and regarding the influence that they may (even involuntarily) exert on lay members.

Chapter 4 presents the second study, in which mock jury experiments were carried out with ten groups of participants, equally distributed between two conditions, each mirroring the procedural rules of the British and the Italian system. The chapter, after providing specific indications regarding the method employed, experimental design, procedure, material, etc., offers a detailed description of the different phases through which the analysis was conducted. A qualitative content analysis was conducted of the data mainly collected through audio-/video-recorded deliberations, while inferential statistical analyses were carried out with quantitative data mainly derived from pre-

/post-deliberation questionnaires. Findings of both the qualitative and quantitative strands were conjunctly reported and discussed.

Chapter 5 contains the final discussion of this thesis, with findings from both studies summarised and interpreted especially in light of their implications with regard to the effects of the two independent variables under analysis. The chapter analyses the impact of the two variables on three broad areas: jurors' behaviour, jurors' errors, and jury deliberation dynamics. Further general reflections are then provided with regard to the two variables, individually considered. The chapter concludes with a section that addresses the main contributions of the thesis, along with an account of its strengths and limitations, and future research suggestions.

CHAPTER 1

LITERATURE REVIEW

'I believe in the jury system.'

O. J. Simpson

1.1 Introduction

This chapter provides introductory reflections on the body of literature most relevant to this research. It begins with a brief historical overview which addresses the most important phases of the introduction and evolution of the jury trial and emphasises how ancient its roots are and how far back in the past we can trace the reasons that govern the use of the jury. Regarding modern juries, an overview of their functioning is provided with specific emphasis on their different composition (purely lay or mixed) and on whether they need to provide motivated or unmotivated verdicts. Since these two factors constitute the foundation of the present research, reflections are proposed in this respect to show that a gap in the previous literature needs to be filled in order to understand whether and/or to what extent these procedural differences may impact jury deliberations. Finally, a review of the literature on decision-making is provided to set the appropriate framework within which this thesis research was developed. Given the extensive research on the topic, the literature review is selectively tailored to the topics of interest to this thesis: namely the analysis of decision-making addressing specific aspects of jury decision-making in both its individual and collective dimension. The chapter then closes with a reference to the two variables under examination, which leads to an explanation of methodological choices, provided in Chapter 2.

1.2 The Jury Trial: exploring the past to understand the present

The jury trial represents a reality which has long been known within criminal justice systems all around the world. The institution is certainly peculiar. In the context of criminal justice systems that mostly involve the work of professionals, the trial by jury brings together a small number of laypeople, with no knowledge of the law and/or of legal procedure, requiring them to find the “truth” about the disputed facts and eventually determine whether or not a defendant is guilty of a criminal act. Given the peculiarity of this institution and the intuitively evident issues that it can raise, it would be difficult to justify the long-lived and widespread use of the jury trial, without referring to the history of its introduction as well as to the core principles from which it arose and that still support its use within criminal justice systems. Indeed, by looking at the foundations that prompted the introduction and wide diffusion of the trial by jury, it is possible to find the reasons that may justify its ability to still be in force, notwithstanding the strong criticisms and doubts that have frequently arisen around its questionable nature and its often flawed functioning.

1.2.1 The origins of the jury: the introduction of the jury in ancient Greece and Rome

As Forsyth (1852) pointed out, when attempting to trace the origin of an institution, it is important to take into consideration the aspects that characterised it when it first appeared, placing it into the context that gave rise to it. With regard to the jury trial, primitive forms of judicial process that placed power into the hands of the citizens may be already found in ancient Greece and Rome. The Greeks, in the shift from aristocracy to democracy, had justice as a main concern. They strongly believed that the most important duty of the State was to ensure that justice could be exerted by its members; therefore, the popular assembly was considered to be the most important democratic body (Stolfi, 2006). Accordingly, popular courts in the fifth century (dicasteries) were formed through a selection of six thousand laypeople, chosen annually by lot amongst citizens of over thirty years of age, who were all eligible to hold the office of jurors (*dikastai*) (Durant, 2011). The selected citizens had to then take the so-called “*heliastic*” oath, with which they committed themselves to decide

on the case in compliance with the law in the first place and, in case there was no law which regulated the specific matter, in accordance to the most “reasonable” opinion. Moreover, they had to pledge themselves to consider only the facts of the case and to listen without any bias to both parties (Hansen, 2003). Besides some differences (e.g., there was no discussion amongst the jurors before they expressed their vote), those “dicasteries” strongly resemble modern juries, to the extent that the principles which lie at the heart of the current institution can be surprisingly found even in this rudimentary form of judicial bodies.

Nonetheless, some would argue that in fact it is not appropriate to trace back the origins of today’s juries to the Greek *heliaea* (Vico, 1904). It would instead be more plausible to identify as an ancestor of the modern jury the popular tribunal set in Rome during the Republic. In the switchover from the regal to the republican period, Rome saw the introduction of two popular assemblies (the Curiate Assembly – *comitia curiata* – and the Centuriate Assembly – *comitia centuriata*) and of the democratic instrument of the *provocatio ad populum*. According to this principle, a citizen, who had been subjected to the *coercitio* (coercive power) of a magistrate, could invoke the so-called *ius provocationis* (namely, the right of appeal to the assembly) against the sentence and obtain a trial before the *comitia* (Staveley, 1954). The *provocatio ad populum* already showed a tendency, which is common to modern legal systems, to detract power from the hands of administrative organs and place it into the hands of citizens. However, it was only with the introduction of a further type of trial, the *quaestiones perpetuae*, that actual juries began to operate within the judicial system. As well as in Athens, jurors were selected by lot, decisions were reached without discussion, and votes were secret (Santalucia, 2010). Jurors in Rome were also laypeople with no cognition of legal matters, yet they could make use of expert jurists’ advice (Frier, 1985).

Both the Greek and Roman judicial systems present peculiar characteristics that can certainly be considered an inheritance that modern jury systems have received from the past. From the brief description reported above, it is clear how ancient the roots of the principle of popular sovereignty are. In the next section, reference to the

introduction of the jury in England and in Italy will be made, in order to better understand the roots from which the two systems that this thesis compares grew.

1.2.2 The origins of the jury: the introduction of the jury in England

The first reference to a judicial body similar to a proper jury in England was made in the Constitutions of Clarendon in 1164, according to which, if a crime was committed by someone rich and powerful and nobody would dare appear against him, 'the sheriff [...] should swear twelve lawful men of the neighbourhood or vill [...] and these were to declare the truth thereof according to their conscience' (Forsyth, 1852:195). With the accused having to face the proof of a trial by ordeal (Groot, 1988) and with the involvement of jurors/neighbours as a body of fact-connoisseurs rather than fact-finders, the first British jury presented significant aspects of difference with modern juries. Although there is common agreement amongst jury systems' advocates on the fact that this institution was introduced for the protection of democratic principles (which it still is supposed to protect), and as a fundamental guarantee for the citizens against the arbitrary State power (Myers, 1979; Hans, 2008), when the jury first appeared in England, it was in fact a form of royal inquisition, a proper instrument into the hands of the royal power (Ploscowe, 1935; Gleisser, 1968). Defined as 'a body of neighbours summoned by some public officer to give upon oath a true answer to some question' (Pollock and Maitland, 2010:148) or, even more plainly, 'a royal tool for prying into the affairs of common people' (Gleisser, 1968:35), the first jury in England appeared to represent, more than a judicial body, an hybrid fusion, that mirrored the spirit of the community from which it was selected (Hostettler, 2004).

Although there is no clarity on the historical process that, over two centuries, generated the transition of this type of jury to the modern jury, its evolution has led to the characteristics and functions that it has nowadays. Despite the differences between old and modern British jury, however, consistency can be found in aspects that the British jury has been maintaining constant since its beginning, namely its purely lay composition, and its possibility to render unmotivated verdicts. While instances of democracy have constantly justified, throughout time, the formal choice

to place the decisional power into the hands of laypeople, fundamentally different reasons were at the basis of the acceptance of unmotivated verdicts. Hostettler (2004) pointed out that, when the first juries came in use, the fact that jurors were not required to motivate their verdicts was due, primarily, to the self-informative nature of the factual knowledge on the basis of which the decisions were made. Secondly, the rigorous value of which the jurors' oath was invested also played a crucial role, perhaps as remnant of the ancient reverential respect towards the sacredness of the ordeals, the oath was considered sacrosanct and the decisions made under oath were perceived as unquestionable as such, without any need for inquiring into evidence. Interestingly, notwithstanding the natural disappearance of these reasons, the British trial maintained, throughout time, their consequent characteristics.

1.2.3 The origins of the jury: the introduction of the jury in Italy

In describing the introduction of the trial by jury in the Italian criminal justice system it is not possible to go as far back in history as it has been done above for the British jury system. Neither rudimentary nor very primitive forms of juries were seen in Italy before the introduction of the institution. The reason for that lies in the fact that the actual predecessor of the Italian jury system is the British jury system. Indeed, although no direct connection can be found between the two systems – since the Italian one is directly inspired by the French system – at closer inspection the link appears clear, as the French system itself used the British one as a model (Apa and Cantarini, 2011).

Moving from instances of popular sovereignty which emerged from the French Revolution and using the British system as a model, the late eighteenth-century French judicial system began to consider the creation of an institution that could guarantee the participation of laypeople in the administration of justice (Padoa Schioppa, 1994). The main idea was that the suspects' guilt had to be not only proven but, perhaps more importantly, recognised by the common popular conscience, in the belief that this was an effective means to restrain arbitrary judgements and to confer to those judgements the strength and fairness of a support coming from the opinion of a college of pairs (Apa and Cantarini, 2011).

The same beliefs and needs were felt in Italy where the idea of the jury as a means of safeguarding freedom and inviolable rights of citizens started growing. The increase of these needs went hand in hand with the progressive loss of trust towards the judiciary power, that was thought to be too close to the executive power and, thus, not provided with the necessary degree of independence that would have been required for a correct exercise of the judiciary functions. In the light of these issues, it was believed that laypeople could counterbalance the lack of independence of the judgments and inject in the criminal proceedings a certain degree of experience and popular values with which judges were considered to be unfamiliar, because of their natural propensity toward technicalities (Di Majo, 2014).

The official introduction of the jury in Italy took place in 1848, although the institution started making its first appearance even earlier (late 1790s/early 1800s) in various Italian regions, that independent to one another introduced the so-called “popular juries” (Di Majo, 2014). The panel of peers was composed of laypeople who were in possession of the necessary requisites to be political electors. From the list of political electors, twelve people (“*giudici di fatto*”, that is literally “judges of the facts”, which mirrors the jurors’ role of fact-finders also in the common law systems) were drawn and placed side by side with a professional judge (“*giudice di diritto*”, namely “judge of the law”, that is to be understood as judge who deals with the more technical/juridical aspects of the matter) and a judge of Appeal (“*giudice d’Appello*”). The panel as such peculiarly composed represented a primitive form of what later became the Court of Assizes (“*Corte d’Assise*”), still in use today.

In terms of functioning, the jury worked as a non-permanent judicial body which had to convene, in periodical sessions (quarterly). The final judicial decision (“*sentenza*”) was unappellable and was rendered by the president of the Court on the basis of the verdict of the jury (Binchi, 2011). This last aspect sheds light on the reasons (still valid nowadays) for the peculiarity of the panel’s composition; since the jury had to render the final judicial decision (a formally written legal document), the presence of professional judges on the panel appeared to be necessary. Therefore, the jury was the first and only judicial body in Italy that required a mixed composition.

The jury system in Italy was, and has remained, a peculiar institution overall. However, whilst most of the above-mentioned aspects have changed over time, some of them have continued to characterise it. The particularity of its composition is certainly one of them and it also represents one of the main differences that exist between Italian and British juries. This difference, along with the requirement for Italian juries to provide a motivation for their verdicts will be addressed in more detail in the following sections, as they constitute fundamental aspects that this thesis analysed, in order to understand their potential effects on jury decision-making.

1.3 World jury systems' differences: not only formal distinctions

The introductory historical overview highlighted a crucial similarity that jury systems share, that is juries have been conceived as a means of democratic guarantee. The democratic principle whereby the decision of certain judicial cases has to be referred to a panel of impartial peers, in deference to the concept of a “fair trial”, seems to be – at least formally – guaranteed nearly everywhere (Vidmar, 2000). On the other hand, despite this similarity, jury systems around the world present also fundamental differences. In particular, common law and civil law nations have developed different ways to observe the above-mentioned democratic principle. European juries significantly differ from Anglo-American juries and from one another in several aspects including: composition, amount of jurors' discretion, and rules for rendering decisions (Kaplan, Martín, and Hertel, 2006). Clearly, these different characteristics also result in different social-psychological dynamics in decision making, which previous research has only partially addressed. Therefore, in order to better understand both the differences and their consequences, this section provides an overview of these differences amongst the various systems.

The most evident difference amongst jury systems lies in the classification of these panels as either “pure” juries or “mixed” juries. Pure juries are those in which the members of the panel are all laypersons, whereas mixed juries are those which require the concurrent presence of both laypersons and professional judges within the panel (Kaplan and Martín, 2006). Broadly speaking, it may be said that the former are mostly

used in Anglo-American systems, whilst the latter are the rule in some European systems (Kaplan, Martín, and Hertel, 2006). However, at closer inspection, it appears clear that this is neither a fixed rule, nor a merely formal repartition. First of all, there are indeed European systems that use pure juries (Spain and Russia, for example) (Martín and Kaplan, 2006). Secondly, this repartition, which might seem to refer only to formal aspects of the panel composition, acquires substantial meaning when it is considered in accordance with another crucial different characteristic; the requirement for some juries to explain the motivation for their verdicts (Kaplan and Martín, 2006).

Indeed, some civil law nations have introduced the duty for their juries to provide an explanation for their decisions as an essential and compulsory requisite for the validity of the verdicts (Apa and Cantarini, 2011). The rationale behind the introduction of this requirement can be found in two main aspects: the requisite is conceived as both a check on the competence of juries and a record that creates the grounds on which future appeals can be based (Kaplan, Martín, and Hertel, 2006). It seems intuitively obvious that the presence or absence of such a requirement may alter the entire jury decision-making process. Nonetheless, whether or to what extent this requirement along with the presence of judges on the panel may impact jury decision-making have yet to be proven and are therefore the main aim of the present study. However it can be argued that the first requirement (motivated verdicts) certainly plays a role in conferring substantial meaning to the second one, the (apparently only formal) difference in juries' composition. It is in fact interesting to note how these two elements – requirement of motivated verdicts and presence/absence of professionals on the jury panel – are intrinsically connected and how in some cases the former inevitably influences the latter.

Under the regulations of the Italian Code of Criminal Procedure, Italian juries are composed of eight members, of which six are laypeople (jurors) and two are professionals (judges) (CPP, 2013). Far from being only formal, this difference has in fact substantial nature and is closely linked with the requirement of motivation. This appears clear when considering the main reason that explains the presence of judges on the Italian panel. Unlike other common law systems (the British one, for instance),

wherein the jurors are fact-finders and render the verdict, whilst the judge deals with sentencing; in the Italian system, as well as in other European systems, these procedures are all accomplished by the jury (Martín and Kaplan, 2006). This justifies the presence of the judges on the panels, since laypeople would not be competent to render and write a judicial sentence. Thus, the Italian jury system represents a good example of how the requirement of motivated verdicts influences the jury composition, as it highlights the need for professional judges in a system that requires juries' written motivated verdicts. On the basis of these arguments, both panel composition and requirement for motivation turn out to be aspects of differentiation amongst jury systems, which – far from being superficial – can in fact substantially characterise the jury as a whole.

It is important to notice, however, that this does not happen everywhere. The binary “pure juries – non-motivated verdicts” and “mixed juries – motivated verdicts” do not appear to conform to a fixed rule. There are indeed nations that, despite using pure juries, require the panel to provide a justification for the verdict. Spain and Russia, partially deviating from the general European trend, have fairly recently (respectively, in 1995 and 1993) introduced a pure jury system, in which only laypeople are part of the panel (Martín and Kaplan, 2006). However, in response to controversies due to distrust of pure juries, both nations have also included the requirement for formal verdict justification (Kaplan and Martín, 2006). What makes possible in this case the coexistence of the motivation requisite and the absence of professional judges is the different way in which verdicts have to be justified. In the case of Spanish and Russian pure juries the justification for verdicts is obtained through a different procedure, which does not require the formulation of written reports. More specifically, jurors' justification is acquired through requiring them to provide answers to a list of propositions about the facts of case. By answering the questions list, Spanish and Russian jurors give a motivation for their “guilty”/“not guilty” verdict.

Subsequently, taking into consideration the jury's verdict and answers to the questions list, the trial judge imposes a sentence (Martín and Kaplan, 2006). Accordingly, in these cases, where no written motivated verdicts are required, the lack of legal competence

on the part of the jurors do not seem to prevent them from justifying their choice on their own. However, this does not amount to say that there are no issues whatsoever related to the lack of jurors' legal competence in these two systems. First of all, the judge has to provide the jurors with instructions and psychological research has widely shown that jurors often fail to understand legal instructions (Wiener et al., 1998). Furthermore, the nature of the questions posed may also affect the conducted reasoning and, thus, the verdict. The questions might be not precise enough, complex, ambiguous, or otherwise misleading. Moreover, as empirical research has demonstrated, jurors' bias may be still found in their answers (De la Fuente, De la Fuente and García, 2003). Therefore it seems that, while posing questions to jurors about their motivation (rather than simply asking them to reach a verdict) could prompt a more evidence-driven deliberation (Hastie, Penrod and Pennington, 2013), this nonetheless does not guarantee a greater effectiveness of the jury trial system.

In light of the previous considerations, it emerges that the jury system can be quite different in design such that the main structures found amongst criminal justice systems are three. Firstly, there are pure juries where only lay members decide without being obliged to justify verdicts. Secondly, there are pure juries where the lay members have to provide a justification for their decisions. Thirdly, there are mixed juries where the verdict is rendered by both laypeople and professional judges, with the further requirement on their part to provide a motivated verdict. Interestingly, both the elements present in the last two categories – professionals and/or motivation – appear as factors aimed at exerting greater control over the lay jurors. While this shows clearly the persistent distrust towards juries, it does not seem possible to identify – on the basis of these factors – a better or worse system, given the widespread occurrence of miscarriages of justice ascribed to incorrect jury verdicts (Ma, 1998).

1.4 Are some jury systems “better” than others?

When considering that the use of the trial by jury has consistently been guided by the intention to counteract the arbitrariness of the State power by referring the

administration of justice to citizens (Vidmar, 2000), it might be argued that the involvement of judges – ‘perceived as [...] a part of the state legal machinery’ (Ivković, 2003:95) – would thwart this intention. On the other hand, those nations that have introduced judges on their juries justify this choice with their distrust towards a purely lay jury trial (Kaplan and Martín, 2006). They concur on the crucial importance of the presence of non-professionals within juries (Langbein, 1981), yet they seem to conceive the presence of professionals as a necessary further guarantee against another kind of arbitrariness, which would be inherent in a judicial decision made by individuals who lack any legal competence and knowledge. Both these arguments appear to start from reasonable assumptions and to come to understandable conclusions. However, in the absence of further research that specifically addresses the point, definite conclusion in favour of one or the other cannot be drawn.

Moreover, with regard to the duty of the juries to provide justifications for their decisions, questions arise about whether this factor may make jurors better decision-makers, requiring a more systematic reasoning or forcing them to minimise the influence of bias and to examine critical evidence in depth (Kaplan, Martín, and Hertel, 2006). It seems plausible that the reasons that permeate a choice enter the decision-making process and are likely, as well as other explanation-based decision-making processes (Pennington and Hastie, 1993), to influence deliberations. It could, therefore, be hypothesised that jurors who are required to identify and report the logical/legal reasons for their decision and the connection of those to the evidence at trial would impart a more evidence-driven nature to their deliberations (Hastie, Pennington, and Penrod, 2013). However, definite conclusions cannot be reached in this respect either. To date those questions have not yet received answers, because the literature on jury decision-making has disregarded motivated verdicts as an issue that could exert an influence on the entire decisional process. Nevertheless, as Kaplan, Martín, and Hertel (2006:120) pointed out, although ‘as yet no direct research on these questions has been attempted [...] these are fertile grounds for social-psychological analysis’.

Accordingly, the evident relevance of both the matters highlighted has prompted the interest of the present thesis to be focused on these factors, in order to gain novel relevant insights. Currently, the lack of definitive empirical evidence to determine whether one of those systems is more successful and effective than the other, does not allow argument in favour or against either of the two. Therefore, reasonably assuming that no criminal justice system can be considered perfect, ‘an examination of other systems... can help to illuminate each system, cause us to reflect on how it deals with problems, and hopefully generate ideas about how it can be made better’ (Vidmar, 2000, p.52). A comparative study never benefits only one country or another (as if there were a “worse” and a “better” system), yet through insights into another reality, the comparison creates an exchange from which both countries might benefit.

Subsequently, a first step to take in order to gain a better understanding of the matter overall is to look at how jurors make decisions. In other words, what psychological mechanisms come into play when a small group of people are required to listen to the facts of a criminal case and eventually decide on another person’s fate.

1.5 Jury decision making – a complex and unique task

Jurors face a very difficult task. These laypeople enter courts and are immersed in a reality that is full of complex, technical dynamics; they attend trials that may be distressful; they listen to and have to evaluate evidence that is often intricate, ambiguous and difficult to understand; they have to use that evidence to eventually come to a decision that will affect another person’s life. When considering how hard it is for most people to make an important decision in general, it becomes even clearer how the peculiarity of a jurors’ position may make their choice and the process to make it overwhelming. Given the uniqueness of this task, in order to fully understand it, a number of areas should be explored. It is clearly important to understand decision-making in general. Also, given the situation wherein jurors have to decide, peculiarities of decision-making under uncertainty have to be considered as well. Furthermore, because jury decision-making is both an individual as well as a collective

process, both these dimensions have to be explored. Implications of all these processes converge into the jurors' task and allow us to better understand it.

1.5.1 Decision-making

Before delving into specific aspects and difficulties of jury decision-making, a description of general decision-making functioning is provided. This will entail an account of theoretical and empirical aspects of decision-making as a whole, in an effort to better understand how people (not only jurors) make decisions.

People's daily lives are permeated by decisions. These may vary in scope, from individuals' choices to large groups and societies' choices. They may also range in seriousness and complexity, from very simple choices to the most intricate and thought-provoking ones (Chankong and Haimes, 2008). Regardless of the type and complexity of choices, 'making decisions is like speaking prose – people do it all the time, knowingly or unknowingly' (Kahneman and Tversky, 1984, p.341). Given the high frequency of decision-making occasions, it is not surprising that the mechanisms that govern its functioning have been explored in the context of various disciplines, which fall beyond the realm of social sciences, embracing also mathematical and economical approaches, all aiming to define the processes that lead people throughout the making of their decisions. Accordingly, the theoretical foundations of the present research project can be found – even before considering specific jury decision-making models – among socio-psychological and economic approaches, which have long been considered as well-established explanations of how people make their choices, and which will be analysed after a brief description of the phenomenon of decision-making and the basic phases of its functioning.

Decision-making is indeed a multi-phased process of identification and evaluation of alternatives. The outcome of this process is the selection, on the basis of the decision-maker's preferences, of one of the evaluated alternatives (Beresford and Sloper, 2008). In any decision-making process, the described evaluation is made on the basis of two main categories of factors: factual elements and value elements. Factual elements are those factors that possess a degree of objectivity which renders them

concretely testable and scientifically verifiable; value elements are, by contrast, those that defy any type of objective scientific verification, in that they are exclusively connected to the individual decision-maker's system of values (Chankong and Haimes, 2008). A combination of these elements enters the process that decision-makers undergo, a process that can be better understood if deconstructed and observed in its individual components.

The entire activity of making a decision may be generally seen as a typical example of a problem solving process, which, according to Chankong and Haimes (2008) consists of five steps. The activation of the process is triggered by the decision-maker's perception of the need to alter the current situation (Initiation step). The situation is accordingly analysed in search for needs or objectives to be fulfilled, so that the overall problem is defined (Problem-Formulation step). At this point, appropriate models for analysis are constructed; these are a collection and operationalization of key problem variables along with their relationship and interaction with one another, intended to be assessed together to shed light on available alternative courses of action (System-Modeling step). Following, the alternatives arisen are compared against a set of decision criteria that are used as measures and must be clearly specified. The measurement functions as a yardstick through which the level of fulfilment of set goals or need can be evaluated (Analysis-Evaluation step). Lastly, each emerged alternative is assessed in comparison to the others on the basis of a pre-specified decision rule set to rank the available alternatives, so that the highest ranking option is the one chosen (Implementation and Reevaluation step). If the "winning" option is satisfactory, the process ends here (in which case, it is defined as an "open-loop process"); if the decision-maker is unsatisfied, the information acquired through this first output will be used again to return to the second step, the Problem-Formulation step (in what is defined as a "closed-loop process").

Theoretical explanations for decision-making mainly fall within two categories: normative and descriptive theories (Kahneman and Tversky, 1984). Normative theories (e.g. formal logic, probability theory, and decision theory) address aspects of rationality and logic in decision-making, that is how people should think, reason and

make judgements about decisions. Descriptive theories (e.g. prospect theory), instead, account for how people actually think when making choices, on the basis of their beliefs and preferences (Over, 2004). The former views the decision-maker as a rational thinker who weights and evaluates; the latter perceives how the decision-maker acts in response to impulse, emotions and bias. Both sets of theories are clearly very relevant, not only to decision-making as a whole, but to jury decision-making as well, in that jurors have been shown to think and consequently act/decide in a way that departs from how they should reason and think in order to reach a fair decision. An exhaustive account of all these theories' propositions would fall beyond the scope of this work. However, it is important to highlight that the usual dichotomy between normative and descriptive theoretical approaches should not be intended as if these sets of theories were mutually exclusive (Suhonen, 2007), since it is in fact the tension between the two categories – the 'ought' and the 'is' – that permeates a great deal of studies on judgment and choice (Kahneman and Tversky, 1984).

1.5.2 Decision-making under ignorance and/or uncertainty

In an effort to narrow down the review of literature towards aspects that are most relevant to jury decision-making, all the peculiarities of juries' decision-making task should be considered. One of them lies in the fact that jurors always choose in a state of ignorance and/or uncertainty. Regardless of how compelling or strong the evidence may be, an intrinsic characteristic of criminal trials is that the objective, "actual" truth of the events is almost always unknown, which in turn suggests that jurors will be deciding while unaware of some aspects of the events. These circumstances place the juror in a very different position from a decision-maker who can decide on the basis of certain elements.

Decision-making literature has clearly defined three different states underlying the deciding process: certainty, risk, and ignorance/uncertainty (Whalen and Churchill, 1971). In decision-making under certainty it is assumed that decision-makers know with absolute certainty the elements on which to make a choice, and this awareness will accompany them throughout their evaluation of alternative possibilities before

reaching the best decision. While this is quite straightforward, attention has to be paid to avoid confusing the other two states; in fact previous research has distinctly defined the two (Knight, 2012). Decision-making under risk regards choices which are made without previous knowledge of their outcomes and consequences, but on the basis of a calculation of objective probabilities (e.g. gambling on a flip of a coin). In decision-making under ignorance and/or uncertainty, a decision outcome is again unknown, but the calculation can be made on subjective probabilities, which the decision-maker must infer or estimate (e.g. investing money) (Kahneman and Tversky, 1984; Wu, Zhang, and Gonzalez, 2004).

As is evident from the description of the three decision-making states, they are all defined in light of a projection into the future, whereby the state in which decision-makers decide is determined by the certain, risky, and uncertain nature of the outcome. However, to highlight the pertinence of this literature to the specific situation in which jurors decide, a further clarification is needed. It should not be neglected that the outcome of a choice is intrinsically connected also to the level of certainty of the elements on the basis of which the choice has to be made. Therefore, if the elements are uncertain (as almost always in the case of jurors who did not witness the events), it is logical to conclude that the decision outcome will be uncertain, since that uncertainty has permeated the entire decisional process. Following this reasoning, it is obvious that the greater the number of unknown elements, the higher the risk of an incorrect decision outcome. It is in this light that jurors' decisions fall within the realm of decisions under uncertainty: jurors, making their inferences and estimates on the basis of uncertain elements, will not know whether they will reach the "best" final choice.

Of course, a determination of which is the "best" choice is highly dependent on the context. More specifically, what all decision-makers have in common is that they will want to make a choice that will produce the highest amount of benefits and satisfaction (Higgins, 2000), since these are the parameters on which the quality of decisions is measured (Carneiro et al., 2014). Accordingly, benefits and satisfaction vary depending on the type of choice; for a juror, they will aim to make a choice that,

for example, prevents a sense of guilt or later regret for discovering to having been “wrong” in determining innocence or culpability. In this sense, the “best” choice, for decisions made in conditions of uncertainty, are quite difficult to identify in advance, because under such conditions there is significantly higher possibility or danger that the outcome of the choice will be detrimental or unsatisfying. For example, it is perhaps unlikely that a car buyer, who has all the necessary elements to make an informed decision, will be unsatisfied with the chosen colour of his/her new car. However it is much more likely that a juror, who lacks many necessary elements to make an informed decision, will be unsatisfied with having let free someone who was later found to be actually guilty or *vice versa*.

Clearly, measuring the quality of a choice on the basis of gained satisfaction means subjecting the understanding of the former to the ambiguity of the latter. Satisfaction itself is in fact very difficult to measure for several reasons (Carneiro et al., 2014). Firstly, decision-makers’ satisfaction is related with – but not only – the choice’s outcome: good choices are those wherein outcomes are more beneficial than the available alternatives. However, also the costs faced to reach the outcome relate to the potential level of satisfaction, hence choices are good when costs do not outweigh benefits (Higgins, 2000). Moreover, the cognitive and emotional effort that the entire decision-making process requires plays a role in determining satisfaction (Oliver, 2014; Bailey and Pearson, 1983). Lastly, satisfaction is obviously influenced by expectations: discrepancy between expectations and reality generate a psychological conflict that decreases satisfaction levels (Carneiro et al., 2014).

Accordingly, in decision-making, a good/satisfying decision is one that ‘has high outcome benefits (it is worthwhile) and low outcome costs (it is worth it)’ (Higgins, 2000, p.1217), but also one that encompasses a ‘sum of one's feelings or attitudes toward a variety of factors affecting the situation’ (Bailey and Pearson, 1983, p.531). It is intuitively understandable that one more factor that affects decision-making satisfaction, by investing all the others, is subjectivity: any decision quality is also determined by each decision-maker’s personal parameters of judgement, because each decision-maker will have a different perception of all the above-mentioned

elements (Carneiro et al., 2014). Benefits, costs, emotions, expectations vary depending on the situation, but also on the person's personality and individual conditions.

All these factors constitute ingredients of any decision-making process and, therefore, of jury decision-making as well. Jurors deliberate in a state of uncertainty generated by their ignorance of some decisional elements (actual to-be-judged events), and naturally reflected on their assessments of benefits, costs, emotions and expectations, hence on the perceived quality of their choice and level of satisfaction. In doing so, the objective difficulty of the situation they experience plays a role along with the subjective dimension of each juror's individual situation. Despite being a collective decision-making task, jury deliberations cannot be analysed without preliminary specific reference to the individual component and its impact on the overall task.

1.5.3 Individual (jurors) decision-making – theoretical and empirical perspectives

'The decision of a group is a weighted sum of the decisions of its constituent members' (Marshall, Brown and Radford, 2017, p.637). Preliminary and fundamental to understand jury decision-making is understanding juror decision-making. Accordingly, research has long been focusing on how jurors make their decisions, taking into account that there is a significant difference between the individual juror decision-making process and the jury's group decision-making process. These areas have been investigated from both a theoretical and an empirical point of view, and for the purpose of this study they will both be reviewed in the following sections.

Starting with the analysis of the individual juror decision-making task, several different theoretical models have been elaborated to describe the processes of acquisition and use of trial information that lead jurors towards their decision. These models can be divided into two broad categories: on one side, algebraic and stochastic models, which attempt to explain human behaviour through the adoption of a mathematical/probabilistic approach; on the other side, cognitive explanations, which

explain the variety of mental processes that jurors undergo (Winter and Greene, 2007).

1.5.3.1 Mathematical/Probabilistic Approaches

The first category encompasses different theoretical approaches, all based on the assumption that jurors carry out mental calculations in order to attribute different weight to the various pieces of evidence presented at trial, so that the produced score is used as a means to determine the defendant's guilt or innocence (Winter and Greene, 2007). Anderson's (1968; 1981) Information Integration Theory (IIT), for instance, posits the existence of three phases: in the first phase (valuation), evidence is perceived and classified in a quantifiable form; in the second phase (integration) this psychological valuation flows into an overall judgement; in the third phase (response), this overall judgement is turned into an observable response. The main assumption of the theory is that people's cognitive processes abide by algebraic/mathematical rules, because judgements of complex phenomena would derive as a result of a mathematical calculation, that is the product of a weighted average of the conducted psychological valuations.

Along the same lines is another mathematical/probabilistic theoretical approach: Bayes' Theorem. The theorem's main assumption is based on algebraic calculation conducted with the probabilistic values of constitutive elements of the event (pieces of evidence, in a trial). In this calculation, the numerical probability value of each piece of evidence is neither affected by the probabilistic values of other evidence nor by the overall product, but it is rather individually calculated. As such, decisions are reached through a comparison of the posterior probability, as obtained through the Bayesian calculation, and the numerical values assigned to the respective standard of proof (Hastie, 1994). Given its assumptions, the Bayesian model is, theoretically, very suitable to juror decision-making, wherein an initial probability calculation has to be adjusted following the collection of each subsequent piece of evidence or testimony (Aitken and Stoney, 1991; Faigman and Baglioni, 1988).

Given that the assessment of evidence and its integration into a legally correct decision lies at the core of jury deliberations, probability theory models have been considered particularly suited to such contexts (Simon, Snow and Read, 2004). However, a series of reasons explain the limited applicability of such theoretical models to jury decision-making. The basic assumption of these theories is that decisions are made through abidance by rules of consistent and logical reasoning, in that individual pieces of evidence are evaluated and combined using probability systems (Aitken and Stoney, 1991; Hastie, 1994). This raises two issues. Firstly, it implies that each piece of evidence is an independent entity, which can be evaluated on its own; in fact, in the majority of trials, evidence is fragmentary, inconsistent, ambiguous and, in order for the jurors to make their decision, it has to be considered in the context of all available evidence and testimony (Simon, Snow, and Read, 2004). Overall probative value cannot be the result, individually calculated, of the probative value of each piece of evidence because the latter is also dependent on other pieces of evidence. Secondly, the underlying assumption that jurors can actually logically assess the trial evidence, and therefore accurately process the information required to reach a fair decision, has not found support in recent research conducted in the field (Greene et al., 2002).

Even without considering these issues, the applicability of such theories to jury decision-making is still limited in scope, in that their usefulness refers only to those instances where mathematical/probabilistic aspects are present in the evidence and can be calculated (for example, DNA evidence, but not eye-witness or alibi testimony) (Aitken and Stoney, 1991). Therefore, these theoretical models are not particularly useful to describe how jurors reach decisions, but provide an effective element of comparison in order to check whether jurors' decisions are reached in accordance to rules of 'correct/logical' reasoning, and to identify errors/bias in their decision-making process (Hastie, 1994; Baron 2008).

1.5.3.2 Cognitive Approaches

By contrast, theories falling into the second category have attempted to explain how jurors reach their decision. These are explanation-based approaches, which focus on jurors' cognitive organization/representation of the evidence. More realistically than

their mathematical counterparts, these approaches depict jurors as active decision-makers who (accurately or not) interpret, assess and elaborate the evidence, rather than passively receiving information to individually employ in a mathematical/probabilistic calculation (Winter and Greene, 2007). The most widely accepted among these approaches is the Story Model of jury decision-making (Pennington and Hastie, 1986; 1988; 1991; 1994). Purposely designed to explain the peculiar decision-making activity that jurors undertake, this model posits that the crucial cognitive process at play in jury decision-making is narrative construction (Pennington and Hastie, 1994): jurors make sense of the trial information by constructing a story of what has happened (Pennington and Hastie, 1988).

According to the Story Model, jurors engage in a three-stage cognitive process when reaching a verdict. The first stage is evidence evaluation through story construction: a story is constructed and all subsequent evidence presented will be assessed in the context of the developing narrative. The second stage is jurors' representation of decision alternatives, and the third stage is story classification, wherein jurors reach a decision by judging correspondence between the information acquired and the decision alternative that best fits the developed story (Pennington and Hastie, 1993).

During this cognitive process, individual jurors' story construction is based on three sources of information. In fact, the trial evidence itself, despite being the only legitimately relevant, is not the only source of knowledge for jurors. Pennington and Hastie (1994) identify two more sources: jurors' personal knowledge about events that are similar to trial events, and jurors' expectations about factors that would contribute to creating a complete story (Pennington and Hastie, 1988; Smith, 1993; Smith and Studebaker, 1996). In other words, since jurors do not come to court *tabulae rasae*, but instead with their own knowledge and expectations arising from personal experiences and beliefs, the assessment of the evidence comes as a result of the application of their background knowledge and expectations to the interpretation of the case facts. All the "ingredients" of the story are thus employed to impose a narrative representation of trial information and, once a story has been created, jurors

evaluate the evidence, and its strength and reliability, depending on whether or not it fits the developed narrative (Pennington and Hastie, 1991).

The influential effect of individual jurors' prior knowledge and expectations can then account for the possibility that different jurors create different stories, even if starting from the same trial evidence. According to this model, jurors are storytellers (rather than fact-finders), and different jurors will create different stories, which in turn may lead to different verdicts (Glöckner and Engel, 2013; Winter and Greene, 2007). However, besides the creation of different stories by different jurors, multiple stories construction can be conducted also by each individual juror, who will consider alternative narratives throughout the trial. During the described cognitive process, prior knowledge is perceived by jurors as factual information along with expectations and attitudes that, unavoidably, they bring to trial. The combination of all this information gives rise to contrasting stories, which are alternative versions of what may have happened, and among which one will be eventually chosen (Willmott et al., 2018). The selective process, through which the "best" narrative will be chosen, is based on four factors ("certainty principles"): coverage (how well the story covers/explains crucial pieces of evidence presented), coherence (how consistent the story is), uniqueness, and goodness of fit (Groscup and Tallon, 2009; Willmott et al., 2018). In the presence of these factors, a story is ultimately accepted and a corresponding verdict is reached by a juror (Pennington and Hastie, 1992; 1993).

These theoretical propositions have been validated by empirical data. Pennington and Hastie (1986), in an early study, asked participants to describe their decision-making process for reaching a verdict on a video-taped fictional murder trial, and found that they consistently reported information in a story format. 'Jurors' stories were not simple lists of evidence', they were rather presented as narratives in which the evidence was set in appropriate causal relations (Pennington and Hastie, 1986:252). The authors also found that participants tended to assign greater weight to information that supported their accepted story, which was further supported by the fact that, when elements needed for the chosen story had not been presented as evidence, participants inferred them on the basis of their personal experience,

knowledge and bias, while removing actual pieces of evidence if unrelated to the chosen story.

Moreover, in a follow-up study, Pennington and Hastie (1988) further confirmed jurors' inclination towards a narrative construction approach, showing that participants presented with evidence in a different way – story-based or witness-based approach – naturally tended to rely more on evidence presented in the story format. Specifically, when the prosecution presented evidence in a story format, verdicts resulted in 78% convictions; when the opposite strategy was used, convictions decreased to 31%. Accordingly, it appears that 'the easier it is for jurors to construct a narrative, the more likely they are to render a verdict consistent with that story' (Winter and Greene, 2007:743).

Consistently, later studies demonstrated, the inclination of jurors towards narrative format. This was recorded also in the context of credibility assessment of witness testimony, wherein it was found that using a story format rather than item-by-item evidence presentation eased jurors judgements (Pennington and Hastie, 1992). Further studies were carried out to specifically test the influence of jurors' background experiences, knowledge and expectations on their narrative construction. For instance, Smith (1991) found that, despite being inaccurate, crime prototypes and schemas often entered the process, since jurors included such incorrect information in their story construction, letting it play a role in the determination of their verdict preference.

Besides influencing jurors' verdict decisions, the choice of a story was also shown to affect jurors' memory. Indeed, Pennington and Hastie (1988) found that mock jurors, in a memory recognition test of trial evidence, were more inclined to recognise trial evidence that was consistent with their verdict/story choice than evidence that would have been in conflict with that choice. However, this may also happen due to the so-called 'pre-decisional distortion', a phenomenon whereby jurors interpret new trial evidence depending on whether it fits with the story selected at the time the evidence

is presented, so that if it is inconsistent with the story chosen, that evidence is disregarded (Carlson and Russo, 2001).

A further development of the Story Model, which can better explain jurors' selective approach to trial evidence, is the Anchored Narratives Theory (Wagenaar, van Koppen, and Crombag, 1993). It postulates that jurors proceed in their evaluation by developing a story about the trial facts and then "anchoring" them in some basic, general rules. It is only finding confirmation in these common sense generalizations, which function as further support for the evidence, that case facts can be recognised as plausible (Bex, Prakken, and Verheij, 2006). Both heuristics and anchors can be found not only in the cognitive processes that jurors undertake when deciding, but also in peoples' daily-life choices. This is partially due to the fact that, using the cognitive tool of heuristics is not a particularly demanding mental effort and, in most everyday situations, it proves to be useful (Tversky and Kahneman, 1974). Moreover, that is also due to the fact that when individuals make judgements about others, they tend to activate an egocentric process which is referred to as 'projection' (Hastie and Dawes, 2010). The issue arising here is that, as Hastie and Dawes (2010) highlighted, this complex judgement process allows bias to enter the decisional mechanism and eventually affect the decision.

Bias can easily stem, for example, from the natural tendency of decision makers to be dominated by their first impressions, the so-called "primacy effect", whereby the information acquired earlier in the process tends to be overestimated and overweighed in the final decision (Nickerson, 1998). Similarly, evidence may be given different weight and interpretation because of people's tendency to reason bi-directionally, that is not constructing their interpretation of the trial facts starting from the evidence, but instead proceeding backwards from their hypothesised interpretations to evidence (Simon, Snow and Read, 2004; Glöckner and Engel, 2013). This phenomenon, referred to as "Confirmation Bias", can be theoretically explained by so-called Parallel Constraint Satisfaction models, according to which when decision-makers try to coherently reconstruct situations based on judgements and inferences, they automatically and unconsciously change their perception of the evidence (Holyoak and Simon, 1999; Thagard and Millgram, 1995; Glöckner and Betsch, 2008).

Accordingly, the biased decision-maker's interpretation of the trial facts will lead the decision-making process, so that the information that supports it will be preferred over the conflicting information, which will be underestimated or discarded (Nickerson, 1998; Glöckner and Engel, 2013).

1.5.3.3 Jurors' Errors

In addressing individual jurors' assessment of evidence, it is worth considering that jurors tend to commit errors (Chilton and Henley, 1996). It seems worth emphasising that, in the case of jurors' decisions, these errors in comprehending the situation can be justified, since they face a very difficult task. Criminal cases are often very complex, especially with the increase of complicated scientific evidence issues that juries are not considered competent to resolve (Myers, Reinstein, and Griller, 1999). Also, the instructions that juries are given are characterised by a linguistic complexity that may represent a further obstacle for laypeople (Severance and Loftus, 1999). Furthermore, there is a significant amount of material to be remembered and it has been demonstrated that jurors' memory for evidence is strongly affected by their bias (Pennington and Hastie, 1988), therefore it is reasonable to presume that jurors will not remember (or will not remember correctly) all the elements of a – sometimes very lengthy – trial. Accordingly, the most common errors may be due to the fallibility of human memory, to incorrect evaluation of evidence, ignorance or misunderstanding of the law.

Broadly speaking, memory is an active process through which people reconstruct events, using knowledge of the past, but for purposes of the present (Malloy, Wright and Skagerberg, 2012). However, it seems that an accurate and comprehensive view of the past is as impossible to obtain as an accurate premonition of the future (Malpass, 1996). The results of an early study led Cattell (1895, p.761 cited in Loftus, 1996a, p.28) to conclude that people 'cannot state much better what the weather was a week ago than what it will be a week hence'. Accordingly, due to the fallible and selective nature of human memory, jurors are very likely to not remember some trial facts. This can happen not only because human beings tend to forget, but also because part of the information might have not been processed whatsoever. Since humans are not

information-processing machines, they naturally activate a filtering system through which information is selectively processed, so that they only attend to what they consider the most important aspects and filter out less important or irrelevant facts (Ainsworth, 1998). If this happens in daily life, it seems reasonable to assume that it will happen even more so during a jury trial, where jurors are overloaded with information.

Psychological literature concerned with the study of memory functioning unanimously acknowledges that the recovery of an event is a three-stage process (Ainsworth, 1998). At the acquisition stage, the selection of event information actually encoded occurs. Subsequently, the retention stage is represented by the time occurring between the event and its recollection. Lastly, the moment in which the encoded event information is recalled is defined as the retrieval stage (Loftus, 1996a). As Loftus (1996a) highlighted, memory errors can occur at each stage: information may be missed and not encoded whatsoever at the first stage; also, information may have been acquired accurately and then undergo transformations (be forgotten, misremembered or subjected to interference, e.g. post-event information) during the second stage; lastly, even recovering available information is not easy, therefore, information may become inaccessible at the final stage.

It is possible that, even when jurors' memory is not a cause of errors and evidence is correctly remembered, its assessment is influenced by jurors' comprehension of it. It has been argued that a group of laypeople 'lack the capacity to comprehend highly technical evidence and apply legal standards so as to render decisions in accordance with the law' (Tarr, 2013, p.153). Indeed, jurors, in order to make their judgement, have to rely on sources of information that are complex and difficult to understand (legal aspects, certain types of evidence, etc.). Jurors' background knowledge, beliefs and bias, as well as contributing to the above-mentioned narrative construction, certainly play a role in their perception and evaluation of the evidence. In this sense, the lack of competence leaves ample opportunity for biased views to enter the evaluation process. Issues of this sort regard almost any type of evidence, direct and/or circumstantial, which jurors tend to under-/overvalue (Heller, 2006).

Particularly difficult to correctly understand for lay jurors (and in fact, for judges as well, as Saks and Faigman (2008) highlighted and criticised) is scientific/forensic evidence, in its actual reliability and probative value.

In the last decade, people's lack of forensic/scientific knowledge has started to be gradually filled by television depictions of it (generating the so-called "CSI effect"), which, while on the one hand, has certainly increased public awareness in that respect (Cooley, 2007), on the other hand, has contributed to misunderstandings and unrealistic views of characteristics, role and probative value of scientific/forensic evidence (Houck, 2006; Schweitzer and Saks, 2007; Cole and Dioso-Villa, 2007; 2009). Science, portrayed as infallible in fiction (Cooley, 2007), is therefore attributed inaccurate meanings and value by laypeople who enter the courtroom with expectations regarding this presumed infallibility (Cole and Dioso-Villa, 2007). A good example is provided by the public (hence, also juries) perception of DNA evidence.

Television depictions have fostered viewers' awareness of the power of DNA evidence: the scene of the criminal leaving their DNA at the crime scene, helping police to accurately identify the culprit of that, and potentially other, crime/s is typical. Furthermore, in real life, the successful use of DNA tests to identify perpetrators, also in cases of serial/multiple crimes commission, has played a role in shaping public perception of the power of DNA evidence. However, even if research confirms that DNA evidence is less likely to be misleading than other evidence, limitations to DNA evidence, although rarely mentioned, certainly exist, and they should not be disregarded by the subjects involved in investigation and trial (jurors included) (McCartney and Amankwaa, 2017). Accordingly, McCartney and Amankwaa (2017) pointed out that a more sceptical approach would be needed to take into account these limitations: potential DNA contamination, identification of partial or mixed profiles and other complex issues should not be neglected, in order to avoid that juries 'are seduced by the purity of the science without fully considering the impurity of its application' (Walker and Stockdale 1999, p.149 cited in McCartney, 2012, p.205).

When not referred to scientific evidence, jurors errors may regard other areas of the legal proceedings, giving rise to law-related errors, such as misunderstanding of essence and application of legal principles, e.g. the reasonable doubt standard of proof, which is very often misinterpreted as corresponding to an unattainable 100% certainty (Laudan, 2003). It is clear that, if the whole deliberation process is governed by an incorrect idea about the threshold that has to be abided by, the final decision that follows such a “contaminated” process, cannot be taken in deference to the desired democratic principles.

1.5.3.4 Further reflections

In light of all considered thus far, in the majority of cases, errors in reasoning and judging occur frequently as a result of ‘mistakes that are made at the very beginning of the process, when comprehending the to-be-judged situation’ (Hastie and Dawes, 2010: 163). When trying to understand that, jurors, as noted, are thought to make sense of the situation through narrative construction. This idea finds wide support in both theoretical and empirical literature. More recent studies also provide general support to the Story Model propositions regarding individual juror decision-making processes. For example, Ellison and Munro (2015), in the context of a qualitative investigation of jurors’ understanding of judicial instructions, assessed the decision-making process that jurors undergo to reach a verdict, and also confirmed their attitude towards story construction of trial evidence. However, while the story model constitutes to date the most broadly recognised and empirically supported theoretical foundation in the field of jury decision-making, its application is not beyond criticism.

One of the issues is that there are theoretical aspects of the model that have not yet been empirically tested. For example, no empirical research has been conducted on the above-mentioned “certainty principles” and on their actual aptitude to orientate the choice of one story over another. Moreover, and most importantly, issues arise because the Model does not seem to cross the boundaries of the individual dimension of jury decision-making. As Willmott et al. (2018) pointed out, despite explaining the individual, pre-deliberation mental processes, this model fails to explain whether and how jurors’ decisions are affected by group deliberations and what are the mental

processes that jurors undergo in a post-deliberation phase. It goes without saying that for a fundamentally collective decision-making task, like jury deliberation is, the fact that the Story Model does not fully account for the group deliberation dimension appears problematic. As well as being the sum of individual decisions, jury verdicts are also the result of elaborate and complex group dynamics, which can strongly influence the overall process and therefore deserve attention.

1.5.4 Group (jury) decision-making – theoretical and empirical perspectives

Broadly speaking, group decision-making is defined as ‘a process in which a group of people, called participants, act collectively analyzing a set of variables, considering and evaluating the available alternatives in order to select one or more solutions’ (Carneiro, et al., 2014, p.368). There are several different theoretical approaches that explain group decision-making dynamics from different perspectives: Social Choice Theory and Social Decision Scheme theory deal with investigating how a society or small group combine or aggregate its members and their preferences in a final collective decision (Laughlin, 2011); Social Impact Theory addresses the relationship between group size and decision-making (Bond, 2005); Group Decision-Making Theory (GDMT) identifies characteristics that groups should have in order to be successful in decision-making (Shelton, 2006). These and other models posit theoretical assumptions that govern group decision-making in general. As such, they may be able to explain certain aspects of jury decision-making, given the collaborative nature of these deliberative processes.

In particular, GDMT may constitute a theoretical framework to investigate and explain the jury decision-making process (Shelton, 2006). The theory postulates that successful decision-making groups present five characteristics. *Small size* is required: the number of people composing the group has to be small enough (six-twelve) for each member to have a chance to interact with the others. A common *purpose* is also necessary, so that the success of each member corresponds to the group as a whole. *Identification* and sense of belonging/membership of each participant also enhances the group

“quality”. Moreover, pace for oral *interaction* is required. Lastly, the adopted *behaviour* has to be based on rules that are accepted by all members (Fisher, 1974). Accordingly, a successful group is ‘a small number of people [...] committed to a common purpose, performance goals, and approach for which they hold themselves mutually accountable’ (Katzenbach and Smith, 1993, p.45).

In an attempt to apply by extension GDMT’s postulates to juries, it has to be observed that juries, theoretically, possess all the above-mentioned characteristics. However, they should be individually considered, as their application to juries may not always be straightforward. Juries appear to present the right trait to be successful in terms of size: juries in most courts are composed of six-twelve members (Shelton, 2006). A common purpose is also certainly identifiable in jurors; they have to assess trial information and apply legal rules to that information, to then reach a verdict (Diamond, Casper and Ostergren, 1989). Issues might arise regarding identification, in that a sense of belonging to the group is difficult to achieve given that this group of randomly assembled strangers, who are required to not discuss the case outside the deliberation room, have during the trial very little occasion for bonding (Shelton, 2006). Likewise, interaction can be problematic, in that many aspects of the deliberation impact on the level of participation of jurors, e.g. jurors’ characteristics/personalities; verdict-/evidence-driven approach (that is, the decision to vote first or evaluate evidence first); unanimity/quorum requirement, etc. Consequently, it is difficult to ascertain that jurors have the same level of participation, as it would be beneficial for a successful decision-making group. Lastly, regarding the adoption of a common accepted behaviour, this appears highly dependent on how group members enter deliberation: group members, who are better instructed on the characteristics of the task, enter the deliberation process with clearer ideas on how to proceed and behave, which currently does not happen with juries (Shelton, 2006).

When investigating jury decision-making through the lens of GDMT, it has to be concluded that juries do not present all the features that allegedly create a successful decision-making group. However, regardless of the specific characteristics, what juries as a social group have in common with other groups is the main foundation of GDMT:

groups are considered better judges than individuals (Hinsz, 1990), which in turn explains (at least, theoretically) why criminal justice systems rely on these small groups of individuals to make such important decisions (Kuhn, Weinstock and Flaton 1994). Indeed, at times, group decisions are of high quality. Groups, through the exchange of ideas, can elaborate solutions that none of their members would have found by themselves; group members may perform better because they are more highly motivated when deciding collectively; groups can also reach a higher creativity level than individuals would on their own (De Dreu, Nijstad and van Knippenberg, 2008). Nevertheless, that is not always the case. Due to particular dynamics, triggered by the peculiar conditions in which juries decide, it is not possible to definitively conclude that group decisions are always better than individual decisions.

As already observed throughout the analysis of the individual dimension of jurors decision-making, juries are composed of a number of strangers who are gathered into the deliberation room after hearing complex, confusing and overwhelming information on the basis of which they have to reach a unanimous decision that holds the moral weight of determining a person's fate. It goes without saying that the described situation is not quite like that of a group of managers meeting at the work place; indeed 'juries are a unique type of social group with strong, bidirectional influence processes in which the individual juror influences the group and the jury influences the individual' (Pigott and Foley, 1995, p.101). In light of all considered thus far, these dynamics and processes of influence are some of the elements that may prevent juries, as decision-making group, from achieving a "successful" result: a fair verdict.

1.5.4.1 Leadership

Group influence may well be generated by dynamics of leadership. Leadership has been defined by early literature on the topic as 'a relationship between one or more individuals and one or more other individuals within the framework of the social unit called a group' (Wolman, 1956, p.11). Contemporary literature has provided a more complex definition, whereby leadership is 'a power-laden, value-based and ethically driven relationship between leaders and followers [...]' (Gini and Green, 2013, p.5). One part of the group is defined as leader/s and another part as follower/s, and the

latter's activities and choices are initiated, stimulated and at times even determined by the former.

Theoretical and empirical literature has addressed the topic with the aim to elaborate and apply leadership theories. Early theories focused on the individualistic dimension of leadership, attempting to identify the characteristics of a leader: Great Man Theory, Trait Theory, Skill Theory, Behavioural Theory all posited that leaders possess traits that differentiate them from followers, whether they are personality traits, given attributes, abilities, actions/behaviours (Harrison, 2018). Typical empirical attempts made to test these theories focussed on identifying leaders and followers within groups and looking for differences, or on identifying patterns in leaders' behaviours. However, all the tested factors were not found to be universal predictors of leadership (Chemers, 2000).

The partial inappropriateness of the above-mentioned models to explain leadership led scholars to move to different approaches which started considering an interaction of individual traits with situational factors, a so-called Contingency Approach, which suggests that leadership effectiveness also depends on whether and to what extent the leader personality fits a given situation (Fielder, 2006). Empirical tests in this case encompassed variables that could reflect how the entire situation affected leadership dynamics. In a similar vein, also the Normative Decision Theory (Vroom and Yetton, 1973) proposed a model that integrated situational parameters into the analysis of leadership, so that the two theories had many features in common and overcame some of the weaknesses of the earlier approaches.

Moreover, in an attempt to further integrate traditional theoretical approaches with factors of relevance to the study of leadership, House and Mitchell (1975) elaborated the Path-Goal Theory, which suggested that a leader's principal purpose is to motivate followers by aiding them to see how their performance can lead them to the accomplishment of their goals. Consequently, leaders' directiveness and/or supportiveness might impact on followers' motivation and performance. Empirically, numerous attempts have been made to test the theory and were based on the

assumption – confirmed by the findings – that a leader's supportive behaviour generates positive followers' reactions and higher motivation and that both specific characteristics of the task as well as of the followers have an impact on the accomplishment of the goals (Chemers, 2000).

The pioneering work on individual, situational, and (directive/supportive) behavioural aspects of leadership certainly continue to provide the foundations for an understanding of leadership dynamics. Important, and greatly relevant to this thesis, is also the subsequent emergence of further approaches which, moving beyond the nature of leadership dynamics, focussed on leadership perception. First studies conducted to investigate followers' perception of leaders found a followers' tendency to evaluate leaders on the basis of the successful/unsuccessful performance, rather than on their personal skills and behaviours (Staw, 1975). Subsequently, other studies found that followers tended to attribute to leaders the causality of certain events, despite the lack of any support for such a connection (Chong and Wolf, 2010). This highlighted the existence of followers' biased views.

According to a follower-centric theoretical approach, leadership was defined through reference to followers' mental processes, and its perception was found to be largely influenced by followers' prototypical views of leaders, which frequently departed from reality. Meindl (1995) identified this phenomenon as the social-constructionist foundation of the so-called Romance of Leadership. Later research, focussing the analysis on individual and group dimensions of leadership perception, also confirmed the existence of followers' biased views. For example, Giessner, van Knippenberg and Sleebos (2008) found that group prototypical leaders (leaders presenting characteristics and behaviours that confirmed followers' cognitive schemas of typical leaders) were considered more effective than non-prototypical leaders.

The presented theoretical and empirical literature, far from being exhaustive on the topic, serves the function of addressing important aspects of leadership and leadership perception, which are certainly of relevance to this thesis. Indeed, social groups are, broadly speaking, natural generators of leadership dynamics; and juries, more

specifically, represent a unique type of social group, wherein bidirectional influences come into play, with the individual influencing the group and *vice versa* (Pigott and Foley, 1995). The very nature of the jury task contributes significantly to the emergence and development within such a group of leadership dynamics. That is, a number of strangers, with different personalities, background and experiences, spend several hours (deliberations may last months) discussing – with the aim to reach an agreement – very serious, thought-provoking matters, characterised by uncertainty and by the burden of the fact that somebody's life is at stake. Such circumstances also make jury deliberations an obvious scenario for the insurgence of contrasting opinions, whether whole-heartedly expressed or silently retained, all intended to pursue (jurors' ideas of) justice.

1.5.4.2 Conformity

The need for jurors' (contrasting) opinions to lead to a unanimous verdict makes further group dynamics come into play. Since it is rare that twelve people agree on a decision at the outset of their discussion, jury deliberation has been defined as 'a study in persuasion and social influence' (Kassin and Wrightsman, 1988: 174). Consequently, previous theoretical and empirical literature have also dealt with dynamics of conformity as an unavoidable phenomenon occurring in group decisions. Social psychological research has compellingly demonstrated, with regard to group interactions, that group members tend to adjust their opinions in order to conform to the predominant and/or most socially acceptable views (Peoples et al., 2012). Conformity is, thus, 'the convergence of individuals' thoughts, feelings, and behavior toward a group norm' (Smith, Mackie and Claypool, 2014, p.315).

Well-established research on conformity has demonstrated that there are various ways in which people may conform. Early models focussed on social influence, distinguishing between Informational and Normative Social Influence, where the former is defined as 'an influence to accept information obtained from another as evidence about reality', and the latter as 'an influence to conform with the positive expectations of another' (Deutsch and Gerard, 1955, p.629). In the first case, influence is accepted due to the individual's intimate desire to be correct; in the second case,

influence occurs in response to need for social acceptance and approval (Devine, 2012). As explained by Smith, Mackie and Claypool (2014), the human tendency of individuals to let other people's opinions guide and influence their own can be prompted by two different, contrasting needs or beliefs. When the group's opinion is perceived as correct, it acquires the power to intimately change the individual's opinion; in such cases private conformity occurs. This means that individuals are actually persuaded that the group's opinion, despite being in contrast with their initial one, is correct and is accordingly accepted as their own. This also generates the consequence that the newly acquired position will be maintained even if the group is not present. By contrast, changes of mind may well be merely external, and be prompted by other circumstantial factors, such as having no choice but to "agree" with the group. This may happen in response to the fear of resulting ridicule or being rejected (Smith, Mackie and Claypool, 2014), or – as in the case of jury verdicts – because it is part of the task to reach a final collective agreement (Foss, 1976). As a result, individuals in this case react to actual or imagined pressure and behave in accordance with opinions/norms that they do not intimately perceive as correct. Contrary to private conformity, when individuals publicly conform, they put forward a superficial change of opinion, to which they will no longer adhere as soon as the source of the actual or imagined pressure (i.e. the group itself and/or the rule, as in the case of the unanimity requirement) disappears.

All addressed so far has been confirmed throughout time by a number of empirical studies. In an early study, Jenness (1932) used a very simple problem as an experimental instrument to analyse conformity. The experiment involved a sealed bottle containing a number of beans that participants were asked to estimate first individually and then collectively. Following the group estimate, participants were asked whether they changed their mind and results showed that the majority changed their individual estimate in favour of the collective one. With the same intention of demonstrating that conformity occurs in uncertain situations, Sherif (1936) employed a more sophisticated experimental stimulus in exposing participants to the so-called autokinetic effect, whereby, due to visual illusion, a spot of light in a dark room appears to move when, in fact, does not. Participants were asked to estimate how far

the light moved and results showed that, whilst the individual estimates notably varied, in the context of collective estimates (gathered when participants were assembled into small groups), groups tended to conclude with a common opinion, which reflected the majority's estimate within-group.

If the benefit of these studies was to show that conformity indeed occurs, even greater credit can be attributed to later studies, which were concerned with empirically understanding the nature (private or public) of conformity dynamics. The first and most renowned empirical attempt came from Ash's (1951) experiment, in which participants had to undertake a vision test whereby they were asked to match the length of a line to three other comparison lines. Participants were divided in groups, wherein only one subject was an actual participant, while the others were research assistants who acted as participants. The task was designed so that the right answer was quite obvious and, in fact, in the control group (lacking potential pressure to conform), only 1 in 35 answered incorrectly. In the experimental condition the participating research assistants had previously agreed on incorrect answers to give and were always required to answer first (aloud). Findings showed that 75% of participants in the experimental condition answered at least one question incorrectly demonstrating they were willing to conform with the (obviously incorrect) majority.

Subsequent interviews shed light on the nature of conformity as observed in the experiments. Some participants reported that they eventually agreed with the incorrect answer; others reported that they became uncertain of their own perception; the majority of participants admitted they did not actually change their mind, but decided to adhere to the group response to avoid being ridiculed or considered peculiar. This clearly demonstrated the presence of elements of both informative and normative social influence, in that conformity was prompted at times by the individuals' idea of a greater knowledge possessed by the group, and other times by their tendency to fit in, which led them to publicly conform.

Indeed, although public conformity occurs frequently, private conformity also exists in multiple situations. One of the most interesting experiments in the field is the one

conducted by Kassin and Kiechel (1996), in which participants had to type letters on a computer, with the warning to avoid hitting the ALT key, which would have made the computer “crash”. After a minute, the computer crashed and the participant was asked whether they hit the ALT key. When asked to sign a confession, 69% of participants complied. Moreover, when one of the confederates pretended to have witnessed the act of pressing the forbidden key, 94% of the innocent participants signed a confession; also, 54% believed they had actually pressed the ALT key (whilst they had not), and 20% created explanations for their non-existent act. Although the experiment was undertaken to investigate the occurrence of false confessions, its findings are very relevant in terms of conformity as they demonstrate to what extent private conformity can occur.

Regarding conformity, the presented overview of theoretical and empirical literature is not intended to be exhaustive, but strictly related to the aspects of relevance to the present thesis. Indeed, according to Peoples et al. (2012), a place where dynamics of conformity find room to occur is small groups, and it is easy to understand that the tendency of individuals to adjust their opinions to comply with a (small) group’s preference is expected to happen even more so among jurors, whose principal task as decision-makers constantly gravitates around opinions, arguments and persuasion, in an effort to reach a unanimous outcome. Moreover, it cannot be disregarded that, unlike other groups, juries are exposed to multiple sources that could prompt a need for conformity: the presentation of arguments and their potential persuasive power affect jurors’ views all the way through the criminal procedure, that is since they enter court and watch the trial unfold, while prosecutors, experts, eyewitnesses, etc. all provide their opinions and defend their arguments, until the end of deliberation, in which they are actively involved (Burnett and Badzinski, 2000).

1.5.5 Experts’ opinions influence

Mechanisms of leadership, influence, pressure and related compliance responses may be observed also with regard to the effects of experts’ opinions on lay jurors (Hosch, Beck, and McIntyre, 1980). Research has shown that, for instance, when jurors heard

experts' opinions on the limitations of the reliability of eyewitness testimonies, they paid more attention to the eyewitness testimony heard thereafter and did not overly rely on the eyewitness identification in cases wherein eyewitness accuracy rates were low (Loftus, 1980; Wells, Lindsay, and Tausignant, 1980 cited in Hosch, Beck, and McIntyre, 1980).

Broadly speaking, experts' opinions are provided because it is recognised that jurors should have an adequate knowledge of the evidentiary material presented to them. Accordingly, as they are laypeople who may well lack that knowledge, when asked to make judgements on matters that exceed the realm of general/lay experience and understanding, reliance is placed on the parties to provide aid – i.e. experts' opinions – for the jurors to understand complex, technical matters (Kirgis, 2002). Jurors see those opinions as valuable sources of information to which they should conform (Smith, Mackie and Claypool, 2014), and indeed rely on them. Previous literature has, quite reasonably and perhaps unsurprisingly, explained this in view of the fact that the perception of knowledge and competence plays a key role in human interactions (Parrot et al., 2015). In particular, Parrot et al. (2015), who exposed mock jurors to experts' opinions manipulating the degree of knowledge shown by them, did not find the experts' low/high knowledge to produce effects on their credibility. This suggests that, in terms of credibility, jurors may not be able to distinguish more/less knowledgeable experts, in that they assume their competence on the basis of their status.

This uncritical acceptance of experts' opinions on the part of jurors has been at the heart of the debate since the legal community began questioning jurors understanding of expert testimony, given its crucial role in affecting the way jurors would evaluate key evidence at trial (Ivković and Hans, 2003). However, without inappropriately focussing the discussion specifically on experts' testimony, it seems worth emphasising that similar issues may be foreseen, by extension, regarding the influencing opinions of any experts on jurors, including judges who work on jury panels. Those judges' role is that of experts who should aid jurors, clarifying matters that are not easily

understandable for them, and ensuring that jurors are in the best conditions to make the important choice they are required to make (Zambuto, 2016).

Therefore, the fact that jurors tend to uncritically accept expert witness opinions', leads reasonably to believe that they reserve the same blind faith to experts/judges working on their panel. The lack of previous research on this specific aspect did not allow to support the claim with past empirical evidence, yet it offered an opportunity to investigate the matter through the empirical endeavours that this thesis presents (Chapter 3 and 4). Considering that the jury trial is a breeding ground for the proliferation of leadership, conformity and influence dynamics in a situation of uncertainty, reliance of jurors on experts' opinions should not be surprising and could be, as it will be later discussed, turned into a beneficial instrument.

1.5.6 Explanation-based reasoning

When discussing decision-making in any field, one should remember that choices, regardless of their seriousness or implications, are always based on reasons, and that a rational agent is an individual who takes those reasons into account when deciding. The making of decisions and the role of the reasons that lead to choices through explanation-based reasoning have been addressed, under different perspectives, within different disciplines. Particularly relevant are theoretical views proposed in the fields of social psychology and economics.

Broadly speaking, the basic assumption of the Theory of Rational Choice is that decision-makers, in any context, naturally tend to look for reasons and/or construct them, as this represents a solution to solve the conflict they face when they need to make a choice (Shafir, Simonson, and Tversky, 1993). While the original theory only focussed on the analysis of this rational process occurring within a rational agent, more recently developed versions of the theory address further specific aspects. Dietrich and List (2013) proposed a novel Reason-based Theory of Rational Choice, which, starting from the same assumption (an individual's choice is based on his/her motivating reasons), also sheds light on the interaction between choices and underlying reasons, emphasising how, given this interaction, people's choices change when their

motivating reasons change. On these bases, what is important to understand – especially for the purpose of this thesis – is what factors may determine a change in the reasons underlying choices; and more specifically whether those reasons may change when the decision-makers know that they will need to be expressed.

While ‘our inner deliberations are silent arguments conducted within a single self’ (Billig, 1986, p.5), things change when choices have to be justified to others. This depends on the different needs that prompt inner and external justifications. Regarding inner justifications, it has been suggested that they might stem from the need to increase self-esteem; to perceive themselves as coherent rational thinkers; to avoid future regret; etc. (Simonson, 1989). On the other hand, with regard to external justifications, the idea is that people tend to make the choice that is conceived as most likely to be justifiable by others ‘such as superiors [...] or groups to which the decision-makers belong’ (Simonson, 1989, p.159). In addition, as a further confirmation of these assumptions, previous research has shown how the need for accountability to an audience enhances decision-makers’ pre-emptive self-criticism, encouraging more coherent and thorough reasoning (Tetlock, 1985). This, in turn, will generate a decrease in the occurrence of primacy effect, fundamental attribution error, overconfidence in judgment, and so forth (Tetlock, 1985).

It appears immediately clear how these theoretical concepts, developed for decision-making in general, may easily apply to jury decision-making. In that context the most justifiable choice (when expressed) is the one that shows a linkage between evidence presented at trial and final verdict. Following this theoretical approach, it is reasonable to believe that jurors – who belong to a particular group (the jury) – when required to publicly express their reasons, will be inclined to look for the most justifiable ones in that context (i.e. one that matches the evidence interpretation). They would then, presumably, discard those choices that are prompted by the effects of personal beliefs, bias, and first impressions, since those would not be perceived as justifiable. Additionally, this might be true even more so where professional judges work on the jury panel; since they might be seen as ‘superiors’ in those contexts, it could be the case that lay jurors look for explanations that they think are justifiable to them.

It seems, therefore, that the need to provide motivations leads people/jurors to automatically reflect on them, which should not generally be taken for granted, as it has been long known that decision-makers are often unaware of the actual reasons that determined their decisions (Shafir, Simonson and Tversky, 1993). By contrast, self-awareness, acquired through explaining verdict preferences to other jurors, enhances consistency (Pigott and Foley, 1995), and most importantly prompts jurors to reflect on motivating factors, stimulating their need to find legally-acceptable reasons.

1.6 Conclusion

This chapter, in an effort to set the appropriate theoretical framework for this thesis, addressed the main topics of relevance to it. The theoretical and empirical literature analysed, despite not being exhaustive, provides robust and well-established insights into those aspects of decision-making and jury decision making, which certainly constituted the necessary foundations for the development of the present research. Considering that jurors' choice can be seen as both an individual (each juror votes) but also a group choice (the final verdict is the result of all the votes), this chapter has addressed both this individual and collective dimension, to then focus on literature addressing topics of relevance to the two variables under analysis. This has created the theoretical and empirical framework within which the studies conducted for this research are set. Following, the next chapter will provide methodological consideration with regard to the studies.

CHAPTER 2

METHODOLOGICAL CONSIDERATIONS

*'In order to decide, judge;
in order to judge, reason;
in order to reason, decide
what to reason about.'*

Phillip Johnson-Laird First

2.1 Introduction

Through the presented review of the existing body of literature on jury decision-making and related issues, several problematic aspects have emerged and deserve careful consideration. Accordingly, in order to offer novel contributions to the literature, this thesis presents two studies, which together investigate the potential impact of two variables on jury deliberations: the presence/absence of professionals (judges) on juries; and the presence/absence of the requirement for juries to provide a motivation for their verdict decisions. The purpose of this chapter is to discuss some general methodological considerations and to justify the methodological choices that have been made in light of the nature of the data needed to answer the research questions. It will, firstly, refer to the mixed (qualitative and quantitative) nature of the data collected. Through a reference to the research questions, it will then introduce the specific methods (interviews and mock jury simulations) that this thesis has employed across the two studies conducted. Accordingly, while a detailed description of the two studies will be provided in the related chapters (Chapter 3 and 4), the research designs developed to undertake them will be outlined in this chapter, with the main purpose of explaining and justifying the methodological choices.

2.2 The overall methodological approach: mixed methods design

This thesis is composed of two empirical studies that were purposely designed so that each study's datasets, besides having their own value, would also inform the methodology and analysis used in the other study. The effectiveness of this approach has been recognised in the previous literature on research methods that emphasised how 'by mixing the datasets, the researcher provides a better understanding of the problem than if either dataset had been used alone' (Creswell and Plano Clark, 2007, p.7). This clearly depends on the fact that the collection of various types of data (quantitative and qualitative) with regard to the same phenomenon grants a more complex and certainly richer view of the phenomenon itself.

In mixed methods research, this view comes as a result of the combination of the objective measurements produced by quantitative analyses with the more subjective and detailed understandings of social meaning that is typical of qualitative investigations (Creswell and Plano Clark, 2007; Jupp, 2012). Consequently, the integrated application of the two methods allows researchers to take advantage of the benefits of both, as well as to overcome the limitations that affect each of them (Steckler et al., 1992). However, this last observation is only true provided that the combination of the methods actually produce complementary strengths and not, on the contrary, overlapping weaknesses (Johnson and Turner, 2003). Indeed, as it is assumed that different methods do not share the same limitations (Rohner, 1977), the effectiveness of mixed methods relies on the principle whereby the weaknesses of one method are compensated, and in fact neutralised, by the strengths of the other (Jick, 1979; Creswell and Plano Clark, 2007). Therefore, attention has to be paid to the accomplishment of this condition and the first step to take in that direction is to understand the strengths and weaknesses of each method.

Quantitative and qualitative methods differ from one another, first of all, in the types of data produced: quantitative research is empirical research where the data are in the form of numbers while qualitative research is empirical research where the data is

mainly constituted by words, narratives, discourse, etc. (Punch, 1998). While quantitative data are the objective reflection of the occurrence of certain phenomena and, amongst other things, of whether or to what extent these phenomena (variables) are related to one another (Steckler et al., 1992); qualitative data are the result of the observation of phenomena in their natural settings, looked at through the lens of the meanings people bring to them (Denzin and Lincoln, 1994). If the quantitative method is able to produce larger datasets, which can then be statistically analysed, it has been criticised for failing to highlight the difference between the natural and social world, in that it does not uncover the reasons that determine individuals' behaviours and beliefs (Filmer, 1972; Noaks and Wincup, 2004). By contrast, qualitative research, with its focus on specific meanings and interpretation, while providing deeper understanding of the analysed phenomena, does not always create generalizable data, as qualitative data are normally obtained from small samples and partly influenced by the researcher's subjective interpretation (Creswell and Plano Clark, 2011; Bryman, 2012).

The research presented in this thesis made use of both qualitative and quantitative methods, as the advantages of one type were required to overcome the limits of the other. Indeed, the choice of a methodology is always closely related to the nature of the data required, and thus collected, to answer the research questions (Punch, 2000), and therefore it is crucial to ensure that the selection of a mixed methods research design actually reflects the needs generated by the research problem. Consequently, to expand on these points, further justifying the choice of a mixed methods design, the next section provides an overview of the research questions on which this thesis is based.

2.3 Research questions

The research that this thesis describes has been conducted across two countries: England and Italy. This was due to the fact that these two countries differ in the presence/absence of professionals on the jury panel and presence/absence of the requirement for motivated verdicts, allowing a comparison that focussed on those two variables and on the effects that they may produce on jury decision-making in practice. Moreover, the researcher's Italian nationality along with her legal professional

background (having worked as a lawyer in Italy) provided a unique opportunity to compare jury decision-making as a function of the independent variables between these two countries.

The over-arching research question addressed by this thesis is:

- Do formal and procedural differences between jury systems (specifically, British and Italian) substantially affect jury deliberation?

Following on from this main question, further questions and sub-questions were articulated and addressed by the two empirical studies. A detailed outline of all research questions and sub-questions that this thesis addresses is provided below.

Figure 2.1: Study 1 – Research Questions

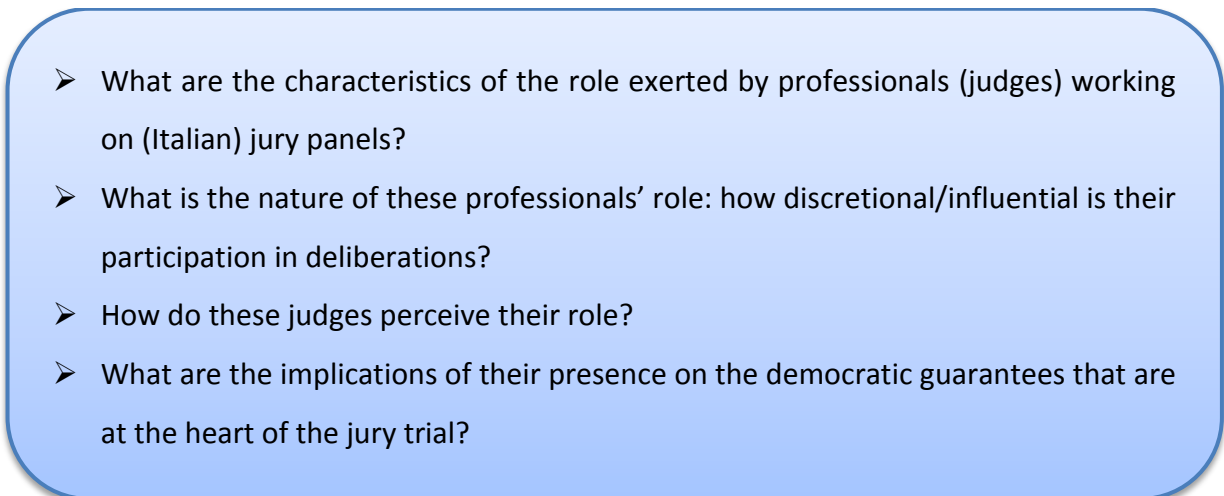
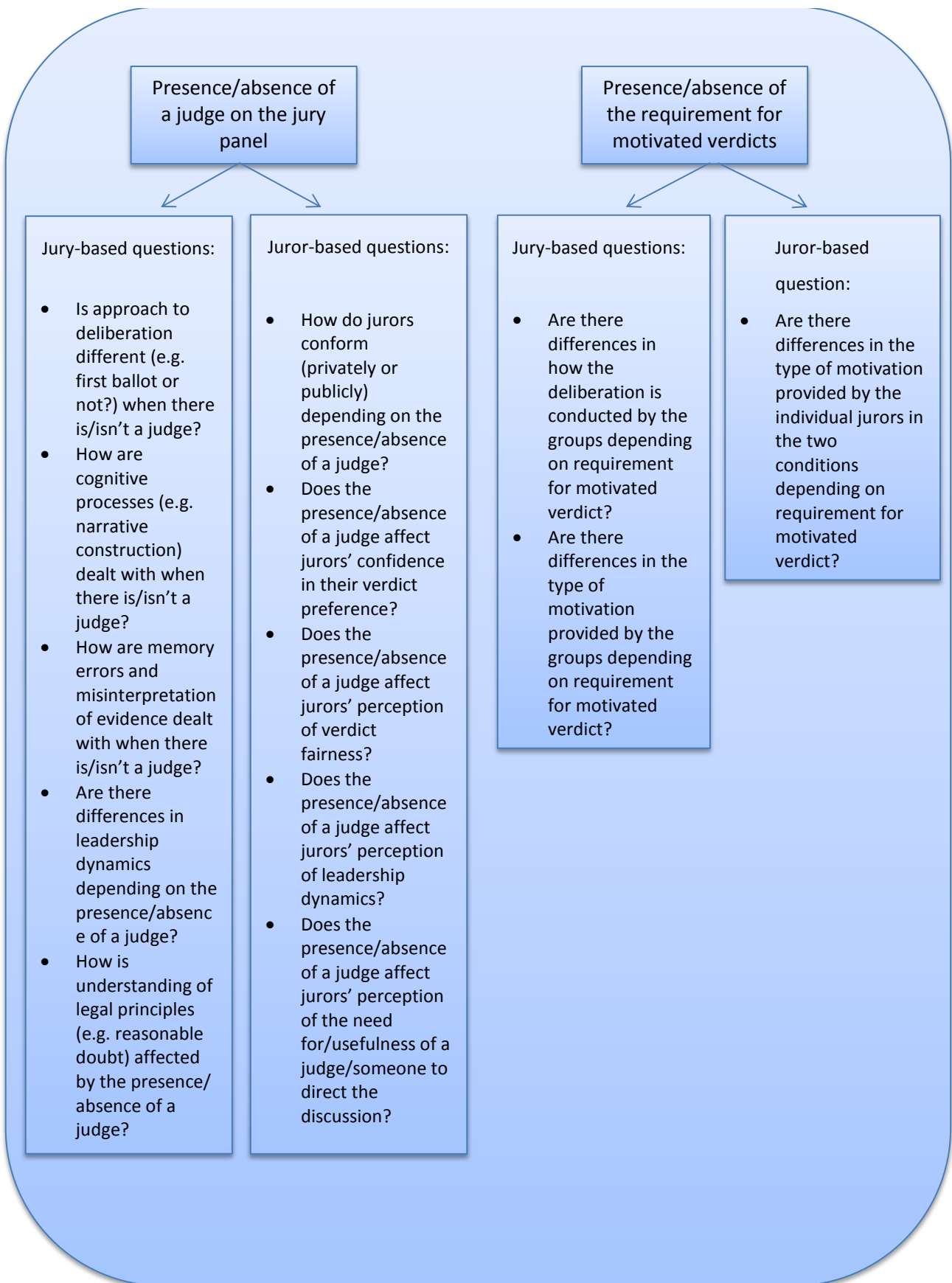
- 
- What are the characteristics of the role exerted by professionals (judges) working on (Italian) jury panels?
 - What is the nature of these professionals' role: how discretionary/influential is their participation in deliberations?
 - How do these judges perceive their role?
 - What are the implications of their presence on the democratic guarantees that are at the heart of the jury trial?

Figure 2.2: Study 2 – Research Questions



As showed in Figure 2.1 and 2.2, questions are divided into categories, depending on the study that attempted to answer them. Study 1 answered questions predominantly related to the role of judges on jury panels; Study 2 addressed two sets of questions: jury-based and juror-based, as it enabled an evaluation of both an individual and a collective dimension of deliberations. A general description of the two studies is provided in the next section to clarify how they were used to answer each of the research questions.

2.4 Research design

Research projects start from theoretical, abstract ideas and are designed based on hypotheses that need to be tested (Draper, 2004). Research designs are, thus, the practical plans of action that create a bridge between the theoretical assumptions and the methods chosen to answer the research questions (Creswell, 2012; Crotty, 1998). Therefore, a description of a research design requires a detailed account of all the practical arrangements that are put in place to gather the needed data (e.g. sample, participants' recruitment, access, data collection, etc.). A detailed description of the specific research designs used will be provided in the appropriate chapters (see Chapter 3 for Study 1, and Chapter 4 for Study 2). However, an indication of the overall research design implemented for this research project will be given here, most of all to emphasise the interconnected nature of the two studies.

2.4.1 Study 1: Italian professional judges interviews

The first study involved interviews with Italian judges who work on jury panels. In the context of the present research project, the idea of gaining those professionals' point of view developed from the impossibility of having access to the deliberation room. In both the UK and Italy access to the deliberation room is prohibited in order to grant freedom to the decision-makers who operate behind its doors. In addition, in the UK – where juries are only composed of lay people – jurors cannot be asked about the deliberation, since prohibitions under the Contempt of Court Act 1981 obstruct research into the content of jury deliberations (Ellison and Munro, 2013). Given these

restrictions, it was considered important to gather Italian judges' views, as the most effective way to obtain first-hand information about jury deliberations in Italy.

To this end, semi-structured, audio-taped interviews with Italian judges were conducted and subsequently thematically analysed. Interviews are considered to be the best method when the scope of the study is to gain a great deal of useful information regarding facts, people's beliefs about those facts, feelings, behaviours, and – perhaps most importantly for this study – standards for behaviours (what people believe should be done in a given situation) as well as reasons behind their actions and feelings (Silverman, 1993). All these aspects were of particular relevance to this thesis, which aimed to gain a better understanding of the judges' behaviour on the jury panel, of their duties and role as well as their perception of this role. Moreover, as it has been compellingly argued by Bleich and Pekkanen (2013), interviews are very useful in studies where they do not constitute the only source of data, but are only part of the data collected in a multi-method research project. In such instances, they are used 'to enhance the internal or external validity of data gathered using other methods' (Lynch, 2013, p.37), which made them particularly suitable for the purposes of this thesis.

The interviews conducted were semi-structured in nature. This method is preferred in qualitative research, as opposed to the structured interviews typical of quantitative studies (Leedy and Ormrod, 2013). In particular, this method is adopted when it is useful to leave a certain degree of freedom in how to cover the topics addressed, so that the discussion starts from a set of defined questions, but it widely expands depending on the interviewee's reaction and on the natural and flexible flow of the discussion. This way the participants are offered a chance to address those topics that they feel are important (Longhurst, 2010), and this, in turn, produces further detailed information to analyse. Participants are also free to spend different amounts of time discussing each topic, which also gives an idea of what they consider more important. In addition, this method has been considered particularly suitable for small-scale studies, while it would be less effective in studies involving a large sample size (Drever, 1995). This element played in favour of this methodological choice since, for reasons

that will be explained in detail later (Chapter 3, section 3.2), this study employed a very small sample (three judges).

Besides the advantages of the use of semi-structured interviews and the appropriateness of this methodology for the present study, it is worth clarifying the pitfalls that interviews might present have not been neglected. Firstly, interviews create discussions that, for the most part, rely on interviewees' memory and it has been widely demonstrated by psychological literature that human memory is often unreliable (Loftus, 1996a), so participants may not remember everything accurately. Secondly, and perhaps more importantly, individuals may be intentionally dishonest, or their accounts might be subtly framed in a way that benefits the participant's world view. Moreover, interviewees might be not very insightful when asked to talk about their behaviours, feelings and motives (Corallo, et al., 2008; Uziel, 2010 cited in Leedy and Ormrod, 2013). Some of these limitations can be, to a certain extent, counteracted by the interviewer probing further during the interview (for example, when participants are being contradictory or not insightful enough), as was done in this study. However, like any study that uses this method, the present investigation cannot be completely immune from the interviews' typical methodological flaws and, given the impossibility of verifying the information provided by the interviewees, it has been acknowledged that those limitations may indeed be influential for this study. However, the consequences of some of these limitations helped unearth contradictions in the interviewees' accounts (when they were being untruthful or not insightful about their roles, for example), which, in itself, was important for the research, considering that it focussed also on the judges' perception of their role. Thus, as the analysis of the interviews demonstrates (see Chapter 3), information gained through this method was, despite the methodological pitfalls, still highly valuable.

The semi-structured interviews were thematically analysed. 'Thematic analysis is a method for identifying, analysing, and reporting patterns (themes) within data' (Braun and Clarke, 2006, p.6). Indeed, given the need to infer knowledge about some aspects of the deliberation process from the insight provided by Italian judges, a method that enabled the researcher to identify, analyse and report themes within data was

required. Since the interviewees were asked all the same questions (as much as possible, given the semi-structured nature of the interviews), the identification of themes was very effective in order to highlight similarities and differences in their opinions on the topics discussed. Indeed, according to Braun and Clarke (2006), a theme is able to capture, within the data, something important for the research question. This is what happened in this instance, because the identification of themes allowed determination of the measure of discretion that judges have in the exercise of their power, as well as several other patterned responses or meanings that are of great importance to answer the research questions that this thesis poses (see Chapter 3 for full results and discussion).

A further consideration supported the use of thematic analysis in this study, as opposed to content analysis, for example. That is, content analysis constitutes another analytic tool apt to identify patterns within qualitative data sets, yet in the majority of the cases it is a more systematic method which provides frequencies and counts and thus becomes suitable for a quantitative analysis of data that were originally qualitative (Ryan and Bernard, 2000). Such a method would have not been appropriate for the analysis of this study's data, given the small sample size and the focus on a few very specific points in the discussion. By contrast, thematic analysis is used with data that produce themes which generally are not quantified. In fact 'the "keyness" of a theme is not necessarily dependent on quantifiable measures – but in terms of whether it captures something important in relation to the overall research question' (Braun and Clarke, 2006, p.10). Following this methodological approach, it can be asserted that, despite the small sample size and the focus on particularly specific topics of interest for discussion, the identification of themes indeed provided important information in relation to the research questions.

This first study's main aim was, thus, to understand the nature of the role of Italian judges, in order to answer the first set of research questions. This aspect was considered crucial for the purpose of the overall thesis, since the presence of professionals is a characteristic of only some judicial systems and since it might seem inconsistent with the democratic principles that govern the jury trial. Whilst these

points will be specifically addressed later, it is worth mentioning here that results from Study 1 played a crucial role in informing the design of Study 2. Indeed, Study 1, by shedding light on Italian judges' behaviours, helped realistically design the mock jury experiments that constituted the heart of Study 2.

2.4.2 Study 2: Mock juries experiments

The design of the second study was made on the basis of what previous literature has determined to be the most appropriate empirical method to investigate jury decision-making, namely the mock jury simulation method (Hastie, Penrod and Pennington, 1983; Ellison and Munro, 2010a). In order to assess the effects of the two independent variables on deliberation dynamics, participants were recruited, divided into groups, and asked to deliberate as a jury on a fictional trial scenario that was presented to them. Following a procedure used in previous research in the field, deliberations were audio- and video-recorded (Ellison and Munro, 2013). Continuing in line with the approach suggested by Ellison and Munro (2013), participants in the study were also required to complete a pre-/post-deliberation questionnaire, which enabled the collection of quantitative data, suitable for statistical analysis. The combination of these two data sources produced a variety of results and data types, which increased the richness and complexity of the results.

In order to take advantage of the benefits of a comparison (Vidmar, 2000) and to find out whether the differences between the two jury systems' (British and Italian) procedures generate any effects on jury deliberations, the mock jury simulations were carried out with groups in two different conditions (one reproducing the British system and one reproducing the Italian system). Experiments with five mock juries in the British condition (that is, mock juries deliberating following the British jury trial procedure) were conducted in England; experiments with five mock juries in the Italian condition (that is, mock juries deliberating following the Italian jury trial procedure) were conducted in Italy. Therefore, participants could not be randomly assigned to a condition; however, this deviation from the gold standard appeared necessary when considering that random assignment of participants across the two different

conditions was not appropriate for this research design. Moreover, the reproduction of British and Italian jury conditions was implemented not only by conducting the experiments in England and Italy, but also by modifying the jury composition and deliberation rules according to the two systems. Therefore, British mock juries were composed of only lay people and had to deliberate without any requirement for producing an official motivated verdict (British condition); Italian mock juries were composed of lay people plus a professional (a lawyer who played the role of a judge) and were required to provide an official motivation for their verdict (Italian condition).

Indeed, it was in designing the Italian condition that the results of Study 1 showed to be crucial not only to answer the first set of research questions, but also to provide experimental instruments apt to increase the realism of Study 2. The greater understanding of the functioning of Italian juries and judges' participation, acquired through the interviews, was a fundamental starting point to recreate, in an experimental setting, mock juries with ecological validity. This was in fact a difficult task to accomplish: a challenge that most previous empirical research employing the mock jury simulation method has not had to face when replicating lay jury conditions. Even without neglecting all the limitations and difficulties that mock jury simulations always present, it is much more straightforward to ask a few lay people to act as jurors than managing and replicating the intricacies and complexities of a mixed jury composition. In the present study, this result was accomplished by having a professional on all the Italian mock juries, who not only had legal knowledge and expertise, but was also instructed on how to play the role of a judge, on the basis of information acquired from the real judges during the interviews.

The effort made in facing this challenge and, even more so, the described experimental strategy adopted to overcome it constitute the elements of novelty of the methodological approach used for this study and the improved validity of the results obtained. Different composition and functioning of the two jury systems were the factors on the basis of which the research hypotheses and questions were designed; therefore, it was necessary that the essence of those differences was reproduced in the mock jury simulations as faithfully as possible. This way, through an analysis of the

pre-/post-deliberation questionnaires as well as of the video-recorded deliberations of the mock juries, the potential effect of the independent variables could be both statistically measured and qualitatively analysed (see Chapter 4 for results and analyses).

The employment of the mock jury method was considered, consistent with the existing literature, to be the most appropriate choice to investigate the variables in this thesis. However, the limitations of mock jury simulations were not neglected. Indeed, criticism against the artificial nature of such simulations has at times resulted in scepticism regarding the contributions that the mock jury paradigm actually makes to the science of jury decision-making (Vidmar, 2008). This has led, at times, to the belief that results of mock jury studies would be impossible to generalise to the real world (Bray and Kerr, 1982; Diamond, 1997; Wiener, Krauss and Lieberman, 2011). Their artificial nature is due to a series of characteristics that all mock jury simulations share and that affect their external validity.

One of the most frequently mentioned drawbacks of mock jury studies is the nature of the sample (Hastie, Penrod and Pennington, 1983; Diamond, 1997). In particular, the frequent use of university students as mock jurors is seen as failing to mimic the composition of real juries. Students are, on average, different from real jurors in several respects: age, education, income, ideology, etc., and these factors are thought to be reasons for corresponding differences in behaviour throughout deliberation (Hastie, Penrod and Pennington, 1983), which in turn would imply inappropriate use of such a sample to accurately represent a realistic jury pool. However, while some would argue that the poor representativeness of the sample may jeopardise the results of studies that use it, research has shown that this might not always be the case. Bornstein (1999) conducted a systematic analysis of jury decision-making experiments, which employed student and non-student samples, in order to find out whether the nature of the sample (and in particular the above-mentioned differences between students and other jury eligible members) had effects on decision-making. He found that in only six out of twenty-six studies an effect of sample on verdict preferences occurred. Moreover, later research has somewhat supported the use of university

students in such experiments, looking at it from different angles. For example, Rose and Ogloff (2001) demonstrated that students can be appropriately used to test the comprehensibility of jury instructions.

In this thesis, the problem of the nature of the sample has arisen, because – as well as in most other studies – university students were used in the present study. The reasons behind this choice are not dissimilar to previous research justifications. When working in an academic environment, students are more easily accessible than members of the general public. Therefore, especially for the British condition (given that the researcher works at a University in England), the five mock juries were mainly composed of university students. This included undergraduate, postgraduate and PhD students, which resulted in some variation in terms of age. An effort was made to obtain an even greater spread of age range in the Italian condition, where – not having direct contact with the university environment – the recruitment had to focus on the general public. Besides opportunity sampling reasons, it is worth pointing out that these choices were also made in consideration of the different minimum age requirements for British and Italian juries (eighteen in the UK, thirty in Italy), which seemed to be respected to a greater extent by having an average higher age range in Italian mock juries.

A further aspect that has been considered as potentially problematic in mock jury simulations regards the materials utilised to create the experimental stimuli through which the mock trial is presented. The use of written transcripts (summarising a hypothetical trial), despite being one of the most widely employed method of mock trials' presentation (Bray and Kerr, 1979; Hastie, Penrod and Pennington, 2002), is also highly criticised for being significantly unrealistic and thus depriving mock jurors of relevant elements they would experience in a real trial. This is alleged to create a lack of interest and involvement in mock jurors that might result in impulsive and/or superficial decisions, leading to unreliable findings (Hastie, Penrod and Pennington, 2002). To overcome these issues, research has proposed various methods for trial presentations, such as live re-enactments, videotaped trials, etc. (DeMatteo and Anumba, 2009); however, none of the proposed experimental instruments has proven

to be able to actually include all the information and experiences that real jurors would go through in a real trial. Moreover, comparisons of the various experimental instruments regarding their impact on mock jurors have shown that, in the majority of the cases, the way in which the mock trial is presented does not have a significant impact on verdict decisions (Wilson and Donnerstein, 1977; Zeisel and Diamond, 1978; Bornstein, 1999). It seems therefore reasonable to conclude that, as well as in any simulations, a certain lack of realism, which mock trial presentations entail, is unavoidable and must be accepted.

Regarding the material used to present the mock trial in this research study, an effort was made to find a balance between the potential effectiveness of the medium and time/costs considerations. Indeed, the use of trial re-enactments or videotaped mock trials would have been expensive and time-consuming and probably not worthwhile given that previous research has reached inconclusive results regarding its effect on jury decision-making (Bornstein, 1999). On the other hand, written transcripts' characteristics appear to be too far from the features of any trial, not resembling what happens in court and presumably fostering mock jurors' disengagement with the matter. In light of these considerations, the material used and the way in which it was presented constituted a middle ground, incorporating a mix of written, oral and visual materials. Chapter 4 provides a detailed description of experimental stimuli and mock trial presentation, suffice it to say here that a hypothetical case scenario was presented by the researcher through the use of a Prezi presentation, containing photos, written testimonies and further pieces of evidence (Appendix F). This way the "story" was presented to mock jurors while going through images and other visual stimuli. The presentation was then concluded with the reading of prosecution and defence closing arguments, which summarised the facts of the case, yet retracing them according to the contrasting views of the prosecution and defence that, like in a real trial, attempted to sway the mock jurors' opinions in their favour.

Broadly speaking, another aspect that differentiates real juries from mock juries, and therefore might be seen as problematic when generalising the findings of such simulations, is the lack of serious consequences as an outcome of mock juries'

decisions. Whilst real jurors are aware that another person's life is in their hands and will take into account the seriousness of their decision since the outset of deliberation, mock jurors know that they are playing a role for the purpose of the experiment and that their choice will not have such repercussions (Ellison and Munro, 2013). However, this lack of consequences is typical of most simulations across several disciplines and it is embedded in the very nature of a simulation, hence accepted from the start when using this method. Secondly, although it is undeniable that the seriousness of jurors' decisions cannot be realistically replicated (as this would be practically and ethically impossible), it has not been conclusively proven that this factor has an influence on verdicts (Hastie, Penrod and Pennington, 1983). Perhaps, even without facing mock jurors with the seriousness of (actual) verdict decisions, it is still possible to sensitise participants so that they take it seriously enough to render a fair (mock) verdict.

That is indeed what appeared to happen in the study presented in this thesis. The impossibility of connecting real repercussions to mock jury verdicts, as a limitation that is inherent in these simulations, could not be avoided or reduced in this study either. Even if actors staged a mock trial, mock jurors always know that they are simply role-playing and that their decisions will not have real implications (Ellison and Munro, 2013). However, there are reasons to believe that, in this as in other studies, this unavoidable circumstance did not affect participants' perception of the seriousness of the task. No lack of interest or engagement was noted during the mock deliberations: participants were evidently concentrating and focused during the presentation of the mock trial and very engaged in the discussion during deliberation. This was proven by the meticulous note taking activity and by the heat and animation of the discussions that the videos of the deliberations show. The mock jurors were expressly asked to take notes, so that the need to undertake that activity would increase engagement. Consequently, the attention paid by them in taking notes confirmed their engagement. Moreover, the behaviour that they adopted during deliberations was left to their own choice, hence the fact that they were animated by the heat of the debate demonstrated their genuine interest and engagement.

Nonetheless, if on the one hand, the mock jurors' engagement and perceivable interest demonstrate attention and involvement in their participation, doubts arise, on the other hand, regarding whether the "voluntariness" of their participation might have played a role, affecting (perhaps increasing) the degree of their interest and engagement, not accurately reflecting what would actually happen with real jurors, whose participation in trials is not left to their own choice. This experimental limitation was considered and its potential impact on the findings cannot be conclusively excluded. However, a few arguments should be proposed to reasonably assess the nature of the issue. First of all, for ethical reasons, participation in research cannot be other than voluntary, therefore this issue could not be overcome and as such is intrinsic to any mock jury simulation. Secondly, it should be considered that, even if real jurors do not voluntarily participate in deliberation, their interest and engagement are presumably naturally generated by the seriousness of the matter and of its consequences (another person's life is at stake). Accordingly, real jurors' interest is stimulated by that element of realism that mock jurors lack; mock jurors' interest might be stimulated by the voluntariness of their participation. Therefore, eventually, regardless of the specific source, interest and engagement are to be presumed in both situations. Lastly, and perhaps most importantly, in numerous occurrences mock jurors in the presented experiments declared to feel anxious and to perceive the pressure connected to the decision they had to make, which would be difficult to ascribe to the voluntariness of their participation alone. These considerations lead one to believe that the issue of voluntary participation and its potential impact on the findings, even if not to be disregarded, were not overly influential for the purposes of this study.

The problematic aspects mentioned here are far from representing a comprehensive list of all methodological issues that the mock jury simulations encounter. They are rather a few of the most relevant and potentially "dangerous" limitations of this method and also those that the present study has somehow attempted to address and – at least partially – overcome. That is not to say that the research design utilised for this study has been able to make any issues vanish; mock jury research limitations are deep-rooted in the method and this study was not immune to them. Nonetheless, to conclude on the methodological choice of a mock jury paradigm, having considered its

limitations and the measures taken in this research to address the main problematic aspects, it is important to emphasise the benefits of the use of the method, since it was in light of them that the design was thought and elaborated in the first place.

The main advantage of using mock juries rather than real juries (if that were possible) is the much higher internal validity of this condition (Hastie, Penrod and Pennington, 1983). The increased internal validity and statistical conclusion are the basis of the choice, made by most existing empirical research to use mock juries' simulations (Wiener, Krauss and Lieberman, 2011). These simulations, in allowing greater control in the experimental conditions, also allow for observation and evaluation of the effects of independent variables much more than anything occurring in natural settings (Hastie et al., 1983). As the study presented in this thesis was indeed aimed at isolating two variables and testing their effects, the paradigm was particularly suitable and was therefore preferred to alternative methods.

2.5 Concurrent and Sequential mixed methods designs

From a methodological point of view these studies were effective in reconfirming the efficacy of the mixed methods design employed for this research, because they enabled the capture of the complexity of the phenomena analysed, through the production of two types of data. This gave a complex and fuller overall picture, as is typical of designs that employ triangulation. Triangulation has been defined as the combination of findings of two or more studies 'conducted to provide a more comprehensive picture of the results than either study could do alone' (Morse, 2003, p.190). Furthermore, as it is an approach that allows the researcher to overcome issues of bias and validity, triangulation is considered one of the most effective ways to improve the reliability of research findings. Moreover, constituting a means for the mutual validation of results, it is particularly appropriate to overcome biases when research is conducted by only one researcher (Blaikie, 1991; Scandura and Williams, 2000). For all these reasons, this approach has proven to be most appropriate for this thesis. This is even more evident when considering that the structure of the overall research design was not strict in the sense of one study yielding one type of data and

another study yielding the other type; in fact, Study 1 produced only qualitative data, while Study 2 produced both quantitative and qualitative data. This allowed the present design to employ two valuable mixed method design options: namely, the concurrent design and the sequential design.

Concurrent mixed methods design has generally been used to validate one type of data with the use of the other type and therefore address different types of research questions (Creswell and Plano Clark, 2007). According to this approach, most times the same participants provide both quantitative and qualitative data (Driscoll et al., 2007). This is indeed what was employed in Study 2, where participants took part in the experiments, completing questionnaires (which were statistically analysed), but were also audio-/video-recorded during their deliberations (which were qualitatively analysed). Sequential mixed methods design, on the other hand, allows an iterative process, whereby data collected at one stage contribute to data collected at the next stage (Driscoll et al., 2007). As noted, this also happened in this research, given that interviews with judges were purposely conducted before the mock jury experiments so that the researcher would be able to incorporate the knowledge acquired during the first phase into the next phase of the investigation. Thus, taking advantage of the benefits of triangulation and of two versions of mixed methods design, this research has collected interconnected data, concurrently and sequentially. It has then analysed them separately and provided a final interpretation that, by looking at the data as a whole, was able to provide a fuller understanding of the phenomena analysed and, in particular, of the impact of the two independent variables on jury decision-making.

2.6 Conclusion

‘Every successful research project requires two things: a meaningful research question and an appropriate way to answer that question’ (Morgan, 2013, p.230). This chapter, through a critical discussion of methodological considerations, has provided an overview of the methodological choices made in this thesis to answer the research questions. This thesis investigated the impact of two independent variables on jury decision-making, namely, the presence of professionals on jury panels and the

requirement of motivated verdicts. Through two interconnected empirical studies and the employment of a mixed methods design, the impact of the above-mentioned variables were quantitatively and qualitatively measured and analysed. Study 1, through the use of interviews with professional judges, yielded qualitative data; Study 2, using mock juries experiments, produced both quantitative and qualitative data. The choice of a mixed methods design (concurrent and sequential) was justified in light of the need, generated by the research questions, to obtain both types of data and to compensate the limitations of one method with the benefits of the other. For the purpose of the proposed methodological considerations, this chapter only provided a general description of the two empirical studies, while later chapters (3 and 4) provide a more detailed account of procedures, materials, participants, specific challenges faced, etc. Accordingly, having explained and justified the methodological choices made, this thesis will now proceed by providing a comprehensive and detailed description of Study 1.

CHAPTER 3

ITALIAN PROFESSIONAL JUDGES INTERVIEWS:

A look inside the deliberation room

*‘To manipulate people effectively,
you need to make everyone believe
that no one manipulates them.’*

John Kenneth Galbraith

3.1 Introduction

One of the most problematic aspects of doing research in the field of jury decision-making is the inaccessibility of crucially relevant information about what actually happens behind the closed doors of the deliberation room. The secrecy of jury deliberation has been granted for most of the jury trial history (Goldstein, 1993) and it is justifiable – indeed, to some extents, also understandable – when considering the negative impact that the fairness of the decision-making process would suffer if jurors did not feel free to express their opinions. In certain countries (such as the UK) this secrecy is not limited to the deliberation phase, but is extended to the further stages; jurors cannot be asked about the proceedings before, during or after the trial (Contempt of Court Act, 1981), and therefore cannot be a source of information about any of the events that usually occur inside the deliberation room, which would be of great value for research in the field. This prohibition has created an obstacle to those research endeavours conducted to investigate jury decision-making processes and which attempt to identify issues that affect the functioning of the trial by jury. Understanding what happens inside the deliberation room, how decisions are made, and what some of the flaws of the system might be, are the first steps to take in order

to find solutions and improve the judicial systems where needed. Accordingly, taking the obstacles into account, novel ways to investigate the issues have to be found.

Given the existence of different types of juries in the various judicial systems worldwide (see Chapter 1, section 1.3) and considering the efficiency of the comparison as a research tool, the study presented in this chapter proposes an alternative way to consider what happens inside the deliberation room, since it cannot be observed directly. The British and the Italian jury systems are compared in this thesis as they adopt different procedures; juries are not required to motivate their verdicts and there are no professionals/judges on the panel in the UK, while instead juries are required to justify verdicts and there are professionals on the panel in Italy. Making use of these differences for the purpose of the investigation, interesting insight was gained by taking a closer look at the Italian system from the inside. Interviews were conducted with Italian judges who work on juries in Italy, in order to obtain greater knowledge about the Italian jury trial from the words of those professionals who work within the system and who could, therefore, offer an accurate account of the decision-making process along with an invaluable opportunity to take a “look inside the deliberation room”.

This chapter will therefore present the conducted study and its findings. The chapter will add specific elements to the general methodological considerations presented in Chapter 2. It will start with a description of the methodology adopted along with an explanation of the methodological choices and an account of the main benefits and limitations that the study presents. Then, a thematic analysis of the conducted semi-structured interviews will be presented. Through an identification of the broad topics addressed and of the themes emerged during the interviews, it was possible to unearth interesting insight into the role of the judges during the various phases of trial and deliberation. Moreover, relevant information about the peculiar aspect of the required motivated verdicts was acquired along with interesting reflections on the management of dynamics of influence within the panel. These findings and their implications will then be discussed and will lead to conclusions on the matter analysed.

3.2 Method

The methodological approach employed for this study was chosen with the aim to learn more about how one of the two judicial systems under analysis – the Italian one – deals with the jury trial, and in particular with the two key factors, presence of judges on juries and need for motivated verdicts. To this end, considering that jury deliberations cannot be attended by subjects who are not involved, there did not seem to be any better way to learn more about motivated verdicts as well as about the Italian judges' role, functions, thoughts and beliefs than to ask these subjects directly.

For reasons that are explained in Chapter 2 (section 2.4.1), the most appropriate research tool for this study was considered to be semi-structured, audio-taped interviews. This investigative instrument, through the flexibility of a semi-structured questions schedule, gave rise to conversations, which unearthed a great deal of useful information regarding the interviewees' role, beliefs, feelings, behaviours, perceptions, etc. (Silverman, 1993; Leedy and Ormrod, 2013). The interviews as such conducted were then thematically analysed. Also in this respect, explanations for the choice of analysis are provided in the context of general methodological considerations (Chapter 2, section 2.4.1). Additionally, in an effort to comprehensively address specific aspects and issues which emerged with the chosen method and analysis technique, further specific methodological reflections and justifications will be presented here.

First and foremost, the sample size requires explanation. Three semi-structured interviews were conducted with Italian judges (two males, one female) who work on jury panels in Italy. To begin with, when focussing on the type of participants required, the reasons for the recruitment of a small sample are intuitively understandable; participants in this study were, given their role and professional position, very difficult to recruit. This was taken into account in the early stages of the research project planning, and yet it was considered worthwhile to try to recruit at least a few judges, given the peculiar and highly valuable contribution that their participation would make to the research project as a whole. The other fundamental consideration made with regard to this stage of the research was that the recruitment of this particular sample, which would have been even more problematic (almost unfeasible) under normal

circumstances, was in this particular case made possible by the fact that the researcher, being also a lawyer (in Italy), had an increased chance of accessing and successfully recruiting judges, through free access to Italian courts and familiarity with the environment and the people working within it.

From the familiarity with the environment came also the awareness that traditional attempts used in research to obtain participants (such as, to call or send out emails) would have been unlikely to generate any responses, as suggested by the literature that highlights the difficulties of recruiting certain categories of professionals to research studies (Bleich and Pekkanen, 2013). As can be easily understood, judges are normally quite busy and tend to pay little attention to those sorts of requests. Therefore, in order to optimise expenditure of time and chances of succeeding, the first recruitment technique adopted required the researcher to contact other lawyers (precisely, ten former colleagues who work in the Italian criminal justice system) to ask them whether they knew judges who possessed the characteristics that participants in this study needed; that is, judges who were working or had worked on jury panels (*"Corte d'Assise"*) in Italy. The employed approach is supported by the previous literature, especially when the required sample is difficult to recruit; 'important actors approached with a referral in hand are more likely to agree to an interview request than those targeted through "cold calls"' (Bleich and Pekkanen, 2013, p.87). The search thus conducted produced six contacts; five judges agreed to have a meeting, during which an overview of the research along with an informed consent form (Appendix A) was provided to them. Among those, three agreed to voluntarily participate in the study and be interviewed.

The reasons for the refusal from the others remain unclear. It is likely due to a desire to maintain secrecy and/or a certain concern about sharing information relating to their job, given the sensitive nature of the cases that are usually tried by juries. Yet, all the judges were expressly assured (verbally and in writing) that they would neither be asked any sensitive questions nor be required to share any information about specific cases – given the nature of the investigation, that type of information was not of interest. Moreover, they were assured confidentiality and expressly told that their

names would not appear anywhere on this thesis and that there would be no way to trace back to them any of their statements. It is unclear why the above-mentioned privacy guarantees were considered to be reassuring enough by some judges and not by others, yet the principle of the voluntary participation was respected and there was no further attempt to convince the judges who refused. The described circumstances confirm the objective difficulty of recruiting such participants.

However, this difficulty is perhaps not surprising, as the category within which the participants in this study fall seems to be the one that previous literature refers to as “elite”, with reference to bureaucrats and/or politicians (Bleich and Pekkanen, 2013). Research demonstrates that these are more difficult to recruit; Temkin (2000), for example, highlights how legal professionals, despite their unique insights/position, are not typically engaged in the research process. Moreover, in the specific instance of the present study, the category of relevance was composed not of judges in general (which perhaps would have been still quite elitist), but of a specific sub-type of judges; the fact that the participants had to be particular types of judges, working in a certain environment and exerting particular functions, narrowed down the possibility to find suitable as well as willing participants. This difficulty was increased by the fact that the judges were purposely selected from different courts, in an effort to gain different views and perspectives. However, despite the awareness that a very small sample might not be ideal for any study, and without any intention to neglect this issue, it was taken into account that qualitative research is mainly concerned with developing rich data that help to explain, understand and shed light on the analysed phenomenon (Moretti et al. 2011). This qualitative study, therefore, even if only based on three participants, reflects this main dynamic/function of qualitative investigations, in that it indeed yielded meaningful elements regarding the aspects analysed.

A few additional observations have to be made about the thematic analysis carried out for this study considering that it was conducted in another country and in another language. These are peculiar circumstances which generate a need for clarification to assure a full and correct understanding of the analysis presented in this chapter. First of all, it is necessary to clarify that the interviews were conducted in Italian, since this

is the participants' and the researcher's first language. The benefits of the researcher and interviewees sharing the same first language are self-evident in the context of a qualitative study wherein the meaning contained in the words spoken are the core of the whole investigation. Being able to identify the subtle nuances of expressions is an advantage from which it is not possible to benefit when, as it might happen with many studies conducted abroad, the researcher's and interviewees' first language is not the same. In such instances they have to pay close attention and, in any case, consider that some expressions may need further clarification (Polkinghorne, 2005).

As Polkinghorne (2005) further pointed out, however, it is also true that translation of data is another factor that might distort meaning. In order to report the results of the present study, though, some data (e.g. interviews extracts) had to be necessarily translated from Italian to English; to make the results available to a wider audience, that data could not be reported otherwise. Consequently, the limitations that are integral to translating qualitative data could not be fully overcome. Nevertheless, it should not be neglected that, given the circumstances (interviewer and interviewees sharing the same first language), the absence of language barriers and translation issues during the data collection stage seems an advantage likely to counterbalance any other downsides. In other words, the possibility for the researcher to fully understand the primary source of the data (Italian native speaking judges/interviewees) has certainly granted a high degree of authenticity to the data during its collection, compared to which the pitfall of the quotes' translation seems to be a surmountable obstacle.

Regardless of the needed translation, it has been asserted that a certain amount of information and its particular nuance is lost, despite any effort, during the transcription, once the oral interview is turned into written text (Polkinghorne, 2005). This is also presented as an insuperable limitation as it is tightly connected to the different nature of oral and written data. Accordingly, in an effort to reduce the negative impact that such a limitation could have on the authenticity of the data, additional attention was paid during the transcription of the audio-taped interviews, in order to retain the majority of those nuances. As suggested by Braun and Clarke

(2006), a thorough and precise transcript includes a verbatim report of as many elements of the speech as possible; both verbal and non-verbal. Therefore, given the advantage that the data was collected by the researcher in person, it was possible to report, during the transcription, all those subtle nuances that were noticed while the interview was occurring. This was possible especially because the interviews were audiotaped; without the need to take notes during the interview, it was possible to listen carefully to the interviewees and to pay attention to their tones and nonverbal behaviour, so that important signals did not go undetected.

Following an approach supported by research methods literature (Bird, 2005), the transcription phase of the present study was intentionally conceived as a first stage of analysis. Therefore, while transcribing, notes were taken to keep track of the researcher's thoughts, reflections, deductions that were prompted by the interviewee's words as well as by some exchanges in the interviewer-interviewee interaction. Such notes, at times, reported thoughts that occurred during the interview and that listening to the audio-tape brought back to mind, and other times consisted of new thoughts and further reflections that were triggered by the possibility of making connections between various statements of the same judge or among them all. This approach has allowed the researcher to start familiarising with the data since the earliest stages of the analysis, so that the transcription did not simply result into a sterile and time-consuming act, yet it became an opportunity to begin to create meanings, in accordance with what represents the overall goal of the qualitative investigation (Lapadat and Lindsay, 1999).

As suggested by research methods literature, the interview schedule was developed taking into account 'the aim of data collection and try to extract data for that purpose' (Elo et al., 2014, p.4). Hence, starting from the research questions (Chapter 2, section 2.3), the interviews (Appendix B) were designed in an attempt to answer them. The use of semi-structured interviews allowed the discussion to start from set points and to then focus on further areas, as they emerged during the interview. The topics addressed are those that correspond to the interview questions, whilst the themes are the patterns of meanings that recurred during the discussion and could be identified

among the interviewees' responses. Moreover, measures were taken to counteract issues that might arise when data is analysed by one researcher. Studies suggest that this is not unusual, and that the credibility of the analysis can still be ensured by making an effort to return multiple times to the data, in order to assure a rigorous and systematic approach (Kyngäs et al., 2011, cited in Elo et al., 2014; Pyett, 2003). This was done for the present study, by rereading and recoding data to ensure that data coded into specific themes, actually reflected the nature of that theme as it emerged across the interviews.

Consequently, the thematic analysis of the interviews will be reported in the following findings/discussion section, starting from the main and broadest topics addressed and then narrowing down towards sub-topics, which in most cases have spontaneously come up as further specifications of the principal topics. However, the themes that emerged while addressing each topic will be the main focus of the analysis. Braun and Clarke (2006) highlighted that simply using the questions posed to the participants as themes would not be appropriate, since in that case there would be no analysis whatsoever. Accordingly, because the ultimate scope of thematic analysis is rather to find repeated patterns of meaning across the data, the participants' responses have been concurrently analysed in order to identify those patterns. The questions posed helped define the topics that were the object of discussion, whilst the themes have been identified later through the reading and rereading of the transcripts. And indeed, as the identification of themes should not be a passive process that results in a simple account of them (Taylor and Ussher, 2001), the themes were identified through an active process of search and analysis, as will be demonstrated later.

To conclude on the methodology, a few considerations about ethics need to be made. To grant compliance of the study with the set ethical standards, a research ethics application was submitted to, and approved by, the University of Leicester Ethics Committee. In order to protect judges' identity and privacy, and to respect the commitment to maintain confidentiality, as granted in the informed consent form they signed, their names will not be reported anywhere within this thesis. Therefore, with the only intent to make their statements distinguishable from one another for the

purposes of the analysis, the three interviewees will be identified as “Judge 1”, “Judge 2” and “Judge 3”. This will be needed in order to compare their responses, verify whether there is consistency among them and for several other purposes of the investigation. Moreover, for the sake of terminological clarity, it should be considered that reference to the Italian lay members of the jury (“giudici popolari”, literally meaning “popular judges”) will be here made by using the word “jurors”, for reasons of convenience.

3.3 Findings and Discussion

In this section the findings of the study will be reported and discussed. An interpretation of them will be given in the light of the previous literature, in an attempt to find plausible reasons behind the judges’ answers, behaviours and choices. The section will be structured as follows: under the following paragraphs each of the topics addressed during the interviews will be reported and, within them, an account of the themes which emerged will be given (see Table 3.1 for a summary of the main topics and themes), with specific reference to the judges’ opinions and thoughts on each matter. In most cases those thoughts and reflections will be expressed through the words of the judges, so that it will be possible to compare the opinions of the three interviewees and, therefore, provide an authentic account of the situation as a whole.

Table 3.1: Italian Judges' Interviews – Topics and Themes

Topics	Themes
Judges' Role On The Jury Panel	<i>Working experience on jury panels</i>
	<i>Training/Professional Development</i>
	<i>Judges' personal interaction with the jurors</i>
	<i>Judges' behaviour/actions prior to the beginning of the trial</i>
	<i>Judges' behaviour/actions at the outset of the deliberation</i>
	<i>Judges' behaviour/actions during the deliberation (how they manage/direct the discussion)</i>
	<i>Judges' behaviour/actions when jurors do not remember relevant facts or evidence</i>
	<i>Judges' behaviour/actions when jurors misinterpret forensic evidence</i>
Motivation	<i>Judges attitude about the need to motivate verdicts</i>
	<i>How do judges manage to write a motivation for a verdict on which they do not agree?</i>
	<i>When is the motivation (reason for choices) asked?</i>
Judges' perception of their role	<i>Influence</i>

3.3.1 Judges' role on the jury panel

One of the broadest and most important topics addressed during the interviews regarded the role exerted by the judges when they work on the jury panel. Understanding what their functions actually are and how these subjects exert them

was important for more than one reason; firstly, because their presence on the panel is one of the main differences between the Italian and the British system, and secondly because the collected information was fundamental for the design of the second study in this thesis. The interviews indeed provided important information about the nature of the judges' role and, especially about the measure of their discretionary power. Since gaining a better grasp on this aspect was the main aim of the entire study, the topic of the role exerted by the judges was one of the most relevant topics investigated throughout the interviews; it came up multiple times, prompting then the emergence of several different themes within it. Below, the themes emerged are listed and an analysis of them is subsequently proposed:

- Working experience on jury panels
- Training/Professional Development
- Judges' personal interaction with the jurors
- Judges' behaviour/actions prior to the beginning of the trial
- Judges' behaviour/actions at the outset of the deliberation
- Judges' behaviour/actions during the deliberation (how they manage/direct the discussion)
- Judges' behaviour/actions when jurors do not remember relevant facts or evidence
- Judges' behaviour/actions when jurors misinterpret forensic evidence

Although the semi-structured interviews did not follow exactly the same schedule, the above-mentioned themes are those that emerged from all the different discussions with the three judges, as recurring identifiable patterns of meaning (Braun and Clarke, 2006), which provided interesting insights into the area of the judges activities, of their actual functions and also of their individual perception of their role from their different perspectives. In the next sub-sections, a detailed analysis of the aforementioned themes is proposed in order to provide a comprehensive overview of the situation as it unfolds in court and inside Italian deliberation rooms.

3.3.1.1 Working experience on jury panels

According to the interview schedule, all the interviews started with the same question about the professional experience of the judges (“*For how long have you been working on jury panels?*”), which was a simple, straightforward question, intended to put the interviewees at ease and to set the scene for further discussion, while also generating useful data (Gill, Stewart, Treasure, and Chadwick, 2008). According to their answers, the three judges had slightly different levels of experience: Judge 1, three years (did not specify when they started); Judge 2, two years (started in October 2013); Judge 3, more than five years (started in November 2009). These slightly different degrees of experience, along with the fact that the judges were purposely selected from three different courts, have helped increase the variety of views and perspectives on the same matters. In fact, if all the judges were, for instance, members of the same court, it would have been likely that they all used the same methods in the exercise of their role and this would have prevented the researcher from accurately drawing out the boundaries of their discretionary power. Similarly, it would have been possible that a similar degree of experience among the judges would have led them to similar choices because of what they had learned until that point.

3.3.1.2 Training/Professional Development

When considering a mixed jury wherein professionals and lay members work together to reach a group decision, an initial interesting aspect about the role of the professionals is how they handle the imbalance deriving from the encounter between their expertise and the other decision-makers’ lack of knowledge of the field. Expectations in this respect relied on the following considerations. The Italian criminal justice system refers the exercise of the judicial power to either a single judge (*giudice monocratico*) or a council of judges (*giudice collegiale*) in other areas of the judiciary, i.e. for crimes that do not require a jury trial. However, in the context of a council, judges work with other professionals with whom they share knowledge and expertise in the legal field. It seems reasonable to believe that, given these underlying conditions, their interaction with each other presents very different characteristics from that occurring between judges and lay jurors. Starting from this assumption, the

question asked was intended to understand whether, given the peculiar nature of that interaction, the judges who work on jury panel receive any particular training/instructions on how to carry out their role. All the interviewed judges replied that no specific training is required in preparation for the execution of their role and they do not receive any instructions. They become judges by the same process as all the other judges in the Italian system, and can then be assigned to that particular type of court (*Corte d'Assise*) that requires juries. The response received on this matter was the same from all the interviewees, therefore this aspect did not require further analysis. However, it was the basis on which the interviews went on focussing on more complex themes.

3.3.1.3 Judges' personal interaction with the jurors

Considering the absence of training, it became interesting to know how the judges determine what the appropriate behaviour is when interacting with the jurors. When asked questions with this regard, they claimed that they simply adhere to the regulations of the Italian Code of Criminal Procedure (*Codice di Procedura Penale – CPP, 2010*). However, when reading the Code, it is clear that it does not provide specific information with regard to that particular aspect. More specifically, in the section related to the actions required from the judges in collegial decisions during the deliberation phase, the Code states (translated from Italian):

All the judges enunciate the reasons of their opinion and vote on each matter separately, regardless of the vote expressed on the other matters. The presiding judge ("presidente") collects all the votes, starting from the youngest judge and is the last who votes. In trials before a jury, the jurors vote first, starting from the youngest (CPP, art. 527, para. 2).

This regulation applies, in general, to all the collegial decisions in criminal justice cases and also, as stated at the end of the paragraph, to jury decisions. As anticipated, however, the law does not give specific indications about how the judges should behave towards the jurors; it only dictates their obligation to comply with a certain voting procedure. Whilst this procedure is clearly intended to grant greater freedom of

expression to the jurors and to reduce mechanisms of influence (as will be seen in greater detail later), it also evidently leaves plenty of room for the judges to in fact independently manage the discussion and their interaction with the lay members of the jury.

Thus, the reference made by the judges to the Code as a “guide” also for their behaviour in the interaction with the jurors in jury deliberations – despite it being actually quite vague in this respect – casts doubts on the accuracy of that response, and might well suggest a lack of immediate awareness on the part of the judges regarding the actual characteristics of their role. In other words, the identification of the Code as a guide comes perhaps as a result of an incorrect and unfounded belief that they are following set rules. A belief that is then spontaneously contradicted when the subjects are naturally led by the discussion to reflect on it. In fact, all the judges (consciously or not), during the course of the interviews, highlighted the independent and discretionary nature of their role both on this particular aspect and in general in the management of the discussion. This is in contradiction with what they stated in the first place; if it were true that they followed the rules of the Code in determining how to behave towards the jurors, then presumably their role could not be discretionary, in that it would simply result from their compliance to the law.

Such conclusion, drawn at the very beginning of the interviews, reveals a reality that has come up repeatedly in the course of the overall analysis; the judges, not used to being asked to reflect on their role, tended to provide their true accounts later, when the discussion became more relaxed and spontaneous. This suggests that there was no intention to lie or hide the truth (which indeed was told later), it was perhaps only necessary to find a way to extrapolate the truth, by prompting them to reflect. Consequently, with the intention to fully understand the nature of judges-jurors interaction, the interviews proceeded with direct questions on this matter. The interaction with the jurors has been generally described by the interviewed judges as quite informal and easy to manage:

Judge 2: *‘Let’s say that, obviously, how to interact with them (the lay jurors) comes rather naturally. We know that we are dealing with people*

who are not professional judges and who have to be somehow instructed on the technical matters... and then everyone gets their own idea.'

The informal nature of the interaction appears to come also as a result of the fact that criminal trials tend to last for a very long time (years, sometimes), during which judges and jurors work together, building a relationship of trust that in turn favours the dynamics of the deliberation:

Judge 1: *'We build a relationship... create a climate of confidence... we deliberate together for a couple of years sometimes! I'm really flexible with them, I become harsh only when they don't show up [...] Even if only one of them is not there, the deliberating session is off. And then you have to explain to them that, no, they can't go at any time and take a flight because their daughter is having a baby in Australia. My colleague and I work with various juries contemporaneously and we can't manage the personal needs of forty jurors... They don't understand this...'*

These observations reveal at least two relevant truths; while lay people – not used to mechanisms of courts and trials – may fail at times to immediately comprehend the importance of their role, the need for their presence proves, on the other hand, how important that role is. Indeed, the real importance of the presence and participation of the jurors was stressed by all the interviewees over multiple occasions during the interviews. They all asserted that the jurors, despite not being professionals, are in most cases level-headed, sensible people, whose participation is highly valued and encouraged. The jurors in those contexts are, according to the interviewees, judges to all intents and purposes.

Judge 1: *'My role isn't that different from the lay jurors' [...] My participation is just as active as theirs; during the deliberation their opinion is as important as ours, their vote counts as much as ours. We are the same.'*

Judge 2: *'They (the lay jurors) are judges just like we are. Their vote counts as much as ours.'*

Judge 3: *'We make them understand the importance of their role, because this is fundamental... because they have to understand that*

they are not marginal in the making of the decisions [...] They are judges to all intents and purposes, just like us.'

If these statements demonstrate that the presence of the jurors is regarded as highly important and influential, further analysis is required with regard to the alleged equivalence of judges and jurors. Although it has been claimed by the judges that jurors and judges are alike, these claims could be misleading when not looking at the bigger picture; at closer inspection, they do seem to require wider interpretation. Whilst all the interviewed judges maintained (and emphasised) that the jurors are like them, they also all acknowledged, throughout the interviews, the existence of some differences in their roles. Once again, as it happened with regard to the faulty perception of the Code of Criminal Procedure as a guide, there appeared to be an initial lack of awareness and consequent discrepancy between what the judges asserted in the first place and what emerged from their words at a later stage of the discussion. In this respect, it is worth noting that, for the purposes of the present analysis, the questions regarding this topic were specifically designed to unearth both the actual characteristics of the judges' role and those resulting rather from their perception of it, as it was taken into account that these two elements might not completely correspond to one another. The existence of differences between the judges and jurors became clearer throughout the interviews when the interviewees explained how they act towards the jurors in the various stages of the trial. Therefore, the analysis of the following emerged themes provides an account of the judges role and functions throughout these various stages.

3.3.1.4 Judges' behaviour/actions prior to the beginning of the trial

In explaining the functioning of criminal trials and jury deliberations, all the judges interviewed made reference to a set of activities that they carry out at the outset of the trial. These entail not only what they do at the beginning of the deliberation phase, but also those behaviours and actions that they adopt in the antecedent phase; before the trial begins, hence even before the jurors sit in the jury box for the first time. The interviewees spent different amounts of time describing this stage, with Judge 3

explaining it in much detail and Judge 1 and 2 giving a more general account of what usually happens on their respective panels.

The judges referred first of all to a preparatory activity, which is intended to familiarise the jurors with the legal/judicial environment, to which they are obviously not accustomed. It can be inferred from the judges' words that this "preparation" is not conducted in a formal manner, it rather happens quite spontaneously, as the judges are faced with a panel of non-professionals. Besides their professional role and expertise, it is easily understandable that, like anyone who has already experienced a certain situation multiple times, the judges would naturally tend to instruct the jurors on what will happen once they walk inside the court and the trial begins. This is indeed what the preparatory/informative activity seems to consist of; a set of information provided to the jurors with regard to the people who will act in court (e.g. prosecution and defence lawyers) and to their roles, to how the trial will unfold step by step, and so forth.

According to the detailed account provided by Judge 3, the following is what happens at the very beginning. Once the jurors are chosen by lot from the list, they are summoned to go to court on the day of the first hearing. The judges acquire the jurors' consent to participate and, after carrying out the formal procedures, begin instructing them. They, first of all, inform the jurors about what is expected of them, on how important their role is, on the type of crime that was committed and will be tried, on how the hearing will be conducted.

Judge 3: *'We need to do that, because they have no idea of what happens in a courtroom.'*

The judges also inform the jurors that, as they most likely lack any legal/judicial notion, guidance will be offered to them during the course of the entire trial and deliberation process with regard to technicalities which will emerge from the words and conduct of the subjects involved in the trial. They thus also explain who these subjects are and what their roles are.

Judge 3: *'[...] a prosecutor, who will explain the facts and show what is his/her plan to get probably to a conviction of the defendant; a defence lawyer, who in turn will show his/her defensive plan; witnesses... you will need to pay close attention to what they say, you will understand from their words how things went and you will have to conduct a particularly thorough exam to evaluate the credibility of those people.'*

The jurors are eventually told more about the evidence, their acquisition, and their relevance to the decision. The judge concluded his description as follows:

Judge 3: *'At this point there is a moment of emotion, because then they realise to have entered a world that's completely different from the one they are used to...'*

Since Judge 3 was the one who gave the most detailed description of how this stage is directed, his description and some of his spoken words were used here to provide a more specific account of what happens in this phase. Although the other two judges did not provide the same amount of detail, their brief and general account about this stage did not contrast with the specific information given by Judge 3. It appears therefore that what is described above is likely to be representative of what happens in most cases. Certainly, the sort of induction that the jurors receive before the beginning of a trial can be seen as beneficial to laypeople who would feel completely lost otherwise. It potentially allows them to overcome the concern often raised with regard to jury instructions in common law countries, given the inaccessibility of the legal language to lay people (Charrow and Charrow, 1979). Nonetheless, despite a greater accessibility of the instructions provided, definite conclusions about the measure of effectiveness of those instructions are difficult to draw. Doubts arise, for instance, on whether urging the jurors to pay attention to witnesses' testimonies would be enough to also obtain from them a fair evaluation of the witnesses' credibility; paying attention does not seem to provide individuals with the necessary instruments to make such assessment (especially in Italian criminal trials, wherein there are criteria for witnesses credibility evaluation – Fadalti, 2008), which is in itself very delicate and, therefore, highly problematic even for professionals.

3.3.1.5 Judges' behaviour/actions at the outset of the deliberation

Once the preparatory phase is completed, the deliberation begins. Albeit brief, mention to the moment of the actual outset of deliberation was made by all the interviewed judges. Indeed, in describing the overall functioning of the deliberation process, they all referred to that initial moment following the passage from the courtroom to the deliberation room. At that stage the jurors will have already come to know the facts of the trial, had access to the evidence and listened to the prosecution's and defence's version of the story. It goes without saying that, for a trial that may last years, there is an overwhelming amount of information to remember, make sense of and coherently evaluate. Presumably, the activity that processing that type of information would require might well turn out to be difficult even for a professional, because of the cognitive effort that it would entail in the first place, and given the fallibility of memory in accurately recalling details (Howe and Knott, 2015). However, the judges have on their side the possession of existing knowledge and previous experience of the field, which has been proven to help understand new material in a given context (King, 1993). It is thus clear how the difficulty increases for a layperson who is in such contexts faced with a significantly more cognitively demanding set of activities (i.e. not only remember, but also familiarise with a most likely unpleasant new environment, understand legal procedures, manage the stressful situation while focussing at the same time, etc.). The judges, perhaps as a means to counteract or at least reduce the confusion deriving from this situation, use a few expedients, which they described during the interviews as follows:

Judge 1: *'To make everything clear, I provide them, first of all, with a summary report which addresses the various aspects of the trial. I also remind them of the specific count of indictment against each defendant... sometimes there is more than one and this can be confusing. I then answer their questions if they want clarifications...'*

Judge 2: *'We study the case file, we study it well and then report to the other members of the panel who have to decide with us [...] Let's say that the report is aseptic (unbiased), as much objective as possible. That is... the facts are described, the pieces of evidence are reported [...] after which everyone gets their own ideas, but, say, the approach is absolutely aseptic.'*

Judge 3: *'I show them what the count of indictment we are going to discuss is and ask them to read the committal for trial [...] They all receive the minutes of the hearings because I tell them: "Look, it's true that I guide you, somehow direct the operations, but you have to be prepared and understand...and please read the testimonies". I demand, then, that each of them reads the testimonies and can concretely formulate an opinion.'*

From the words of the interviewees it appears clear that the discretionary nature of their role and the absence of rules clearly set by the law leaves them with enough space to act as they wish. However, their choices on how to act in this phase turned out to be overwhelmingly similar to one another. When considering, for example, that there is no express indication in the art. 527 CPP of the need to provide a summary report, the fact that two of the judges mentioned it acquires meaning. The reason behind this might be simply common sense, considering how confusing that reality could be for the jurors. Yet, a plausible explanation might be a different one, which emerged in particular during the discussion with Judge 2, who further referred to a summary report as something that is normally used in other collegial decisions:

Judge 2: *'You report to the panel as you would with colleagues, I mean, within a panel there is a person who reports [...] this phase is the same, whether it's done within a panel made of professionals only or within a mixed panel made of both professionals and laypeople... what changes is the technicalities, the legal matters in the strict sense, which obviously don't need to be explained to a professional and instead have to be explained to laypeople.'*

Accordingly, it might well be that the rule of creating a summary report applies to collegial deliberations occurring within panels composed of professionals and, therefore, by extension, used also for jury decisions. This would in this case explain the correspondence of the judges' answers, despite the absence of a set specific rule in this sense, whilst at the same time though it does not contradict the discretionary nature of their role; even if they applied that rule by extension, it would still be their choice to do so.

3.3.1.6 Judges' behaviour/actions during the deliberation (how they manage/direct the discussion)

One of the most interesting phases to investigate in order to fully understand the judges' role within the Italian jury trial is the discussion conducted inside the deliberation room. This is perhaps the stage that is most capable of unearthing procedural/formal and substantial differences between the Italian and the British system. Therefore, a number of questions posed during the interviews were designed to obtain information on this phase; how the judges behave during the deliberations, how they manage and direct the discussion. Also in this case the interviewees described behaviours which, despite not being set by any rule, were found to be quite consistent among each other. First of all, the role was described as directive, of guidance, particularly focussed on clarifying matters and filling gaps of a legal/technical nature.

Judge 1: *'Throughout the deliberation ours is an open and quiet dialogue, during which the jurors receive all the explanations that they request and are free to express their opinions at any time.'*

Judge 2: *'That's our behaviour; purely informative. The only difference that exists within the panel is this: I have greater technical knowledge.'*

Judge 3: *'We try to guide them a little... within the limits of what is possible.'*

The agreement on the general characteristics of the judges' functions during deliberation suggests that in the minds of the judges this is the most appropriate and effective way to behave. And indeed, whilst doubts arose and have been already expressed as for the efficacy of the preparatory phase and the sort of induction that jurors receive, the beneficial effect that the directive and clarifying intervention of professionals during deliberation might have on the jurors' comprehension of the trial dynamics cannot fail to be seen. However, focussing once again on the discretionary nature of the judges' role, it is worth noting how Judge 3 referred to 'limits', as if there were set boundaries, beyond which directing the jurors would be wrong or inappropriate, if not even forbidden. Nonetheless, excluding that those limits are set by the law (as seen previously, they are not), the meaning of that phrase has to be

intended through the assumption that the judge was in fact referring to limits set by self-imposed rules. This hypothesis found confirmation throughout the interviews and is supported by psychological literature, as will be explained at the end of this chapter.

Following with the description of the judges' behaviour during deliberation, other characteristics of this role were emphasised:

Judge 1: *'I play the role of, say, an arbitrator during the discussion, making sure that none of the members of the panel predominate over the others or don't let the others express themselves freely, so that everyone has a concrete opportunity to express their opinion. Dynamics of pressure and attempts of leadership can occur in any discussions, also in panels made up of only professional judges, and they occur even more so in panels that include laypeople [...] It's advisable though that all the members express their opinions and I try to grant them this freedom.'*

Judge 3: *'We listen to them, we do. I mean, there is a participation... but I want it, because I wouldn't want to deal with people who are there just to make up the numbers. Here's what I do: I don't make them speak later, I let them speak first, so that everyone can express their opinions. But I want all of them to express themselves... I see if there is agreement, if there is somewhat harmony among them... because then they'll talk also when we are not there...'*

As the quotes above show, it is the judges' concern also to try and ensure jurors' freedom of expression, so that the occurrence of detrimental dynamics of persuasion, influence and conformity – ineradicable characteristics of group decision-making which thus come into play behind the closed doors of the deliberation room (Lloyd-Bostock, 1988) – can be avoided or, at least, reduced. Once again, it seems reasonable to conclude that this cannot be seen as anything other than a benefit, which would contribute to a fairer discussion. The same goes for the quote below, which is worth mentioning here in analysing the multi-faceted approach that the judges have during the deliberation phase.

Judge 3: *'But I, for example, when they hear at the end of trial the prosecutor who offers a reconstruction... especially when he/she is good... the prosecutor offers a nice reconstruction of the facts, with all the indications and details and so on... Then they say: "oh, that's how it went..." and then I say: "Wait, hold on a second, because when you'll hear the defence lawyer you'll realise that the certainty you have now will be lost".'*

And in fact after hearing the defence lawyer... again, if they are good they demolish each point... they go: "Oh, but it wasn't exactly like that then". And then it is difficult, delicate to make them understand that some passages have been highlighted by the defence, but there are many others that do not fit with what he/she argued.'

This last interview extract was included here as it seemed to raise a crucial point in the on-going analysis. It contains an excellent example of how and to what extent the role of guidance finds application in reality. The described dynamic certainly has significant explanatory value in showing that the directive role of the judges is in fact not restrained to an explanation of the technicalities and an illustration of the intricacies of the trial (which indeed is needed), it also stimulates the critical conscience of the jurors, so that they are more cautious about those elements that, as non-professionals, they might more easily neglect.

Obviously, the way in which this role was depicted was neither the only analysed element nor was it blindly believed as to faithfully represent the facts. It was taken into account that the provided descriptions might merely be the result of the judges' perception of their role and as such differ from what in fact happens. Therefore – given the impossibility to actually watch a deliberation and verify the accuracy of those accounts – cautious analysis was conducted in order to identify contradictory statements that would disprove the accuracy of the description. However, differently from what happened with regard to dynamics of influence (as addressed later, in section 3.3.3.1), no evident contradictions were found in the description of this phase of the discussion.

3.3.1.7 Judges' behaviour/actions when jurors do not remember relevant facts or evidence

Taking into consideration the context of the discussion conducted in the deliberation phase and given the role of guidance and direction that the judges claimed to play because of their legal knowledge and expertise, it seemed plausible that at times their function would also involve keeping the discussion focussed on relevant aspects (e.g. evidence, counts of indictment, etc.) as well as bringing attention to those elements that might have been forgotten. On the basis of these assumptions, the interviewees

were expressly asked questions about this, and their answers confirmed that those tasks are indeed part of their role.

Judge 1: *'I always make sure that the jurors remember the specific aspect we are talking about. I ask them explicitly if they remember it before carrying on with the discussion.'*

Judge 2: *'The assessment of evidence is a technical aspect, let's say, which has to be explained and deepened... How a piece of evidence has to be evaluated, what value it can have, if it's sufficient, if it needs more proofs, if that's circumstantial evidence or "full" evidence... These are all things that have to be told to those who have to decide.'*

Judge 3: *'Sure, surely. This task (reminding the jurors of relevant elements) is essentially done when we are about to decide, because it's then that they are most incisively engaged in the process.'*

Indeed, when imagining a deliberation carried out by a jury composed of only lay members, the fact that jurors might tend to focus on the "incorrect" aspects – i.e. those that are not relevant or evidence that should not be considered because it is inadmissible (London and Nunez, 2000; Steblay et al., 2006) – and forget about the relevant ones, is perhaps one of the most worrying scenarios. In this sense, the contribution that judges may provide in a mixed jury context to ensure that jurors do not forget relevant evidence is evidently beneficial. The contrary might also be true when jurors are required to forget inadmissible evidence. Such pieces of evidence are often presented at trial for two reasons: either because, in the natural unfolding of the trial, it is impossible to have control on all the events – e.g. a witness might provide a prohibited testimony because they do not know the rules of the proceeding – or because lawyers might want to make a strategic move by introducing a piece of evidence that is useful to their case, despite knowing that it will be declared inadmissible later (Demaine, 2008). In both cases, jurors will be asked not to consider the inadmissible evidence for the purposes of rendering their verdict. In pure jury systems the sole instrument that can be used to attempt to achieve this difficult goal is the judges' instructions which urge the jurors to disregard that piece of evidence. As Demaine (2008) highlighted, though, the actual efficiency of those instructions has

been the object of heated debate and doubts remain about the concrete possibility of erasing testimony from jurors' minds simply by asking them to disregard it.

Whether there is the need to make individuals remember, or forget something, a fundamental factor cannot be neglected; human beings have an asymmetric control over their memories, whereby while it cannot be presumed that people (and in this case jurors) will be able to remember everything, it cannot be disregarded that removing certain information from one's memory is difficult and in fact, at times, impossible (Frey, 2005). It is in this light that comparatively considering the situation through the lens of a system that uses mixed juries becomes interesting. Indeed, without claiming that the presence of judges during the discussion will be enough to achieve the goal of making people forget information, it is plausible that it creates more favourable conditions for a fair discussion. Knowing that in a mixed jury system such issues are potentially solved, or at least dealt with, becomes an important element of the comparison proposed in this thesis and will be further discussed when addressing the theme of the potential influence that judges might exert on jurors, in order to make a fair evaluation of both the advantages and disadvantages of the use of a mixed jury system.

3.3.1.8 Judges' behaviour/actions when jurors misinterpret forensic evidence

Related to the theme analysed above is the possible misinterpretation of forensic evidence on the part of the jurors, and the consequent need for them to receive advice in this respect. During the interviews all the judges were asked questions regarding how they generally manage situations wherein forensic evidence is presented at trial and they thus have to face the eventuality that jurors may not fully understand or misinterpret the probative value, as has been demonstrated both in the research literature (Smith, Bull, and Holliday, 2011) and in analysis of miscarriages of justice (Innocence Project, 2018). Those questions started from the same assumptions as the ones referred to above; the likelihood that non-professionals might more easily forget, misremember, misunderstand technical aspects of the trial, and that they might need therefore aid to make a correct evaluation. This appeared to be true even more so for

forensic evidence, which in most instances presents a very high degree of technicality. This was confirmed by the responses received from the interviewed judges, which also provided some additional elements of reflection.

Judge 1: *'Well, that is normally the experts' job, especially because regarding some technical aspects of that sort of evidence we also lack the necessary knowledge to make accurate evaluations. The experts talk in terms of probability and, depending on the type of trial, there is a different degree of required probability... (to admit the evidence)'*

Judge 2: *'The evidence can be read in various ways, but fundamentally one gets an idea... the professional judge as well as the lay juror... so one listens, looks and then is able to catch the nuances... Forensic evidence, like DNA etc. is scientific evidence regarding which we and they (the jurors) have the same approach, I mean, if the expert says something... if the expert, say, concludes with certainty then there is little to discuss. If the expert instead concludes in terms of probability, then obviously that thing has to be evaluated along with others.'*

Not only do these answers confirm that lay jurors might have issues interpreting forensic evidence, they also show that judges themselves are not immune from the possibility of making the same sort of errors. The interviewees admitted to being in the same position as the lay jurors, given that despite being professionals in their field, they are not scientific experts and have no competence in certain matters. Previous literature on jury decision-making has compellingly demonstrated that jurors tend to rely widely on experts' opinions rendered through testimonies during criminal trials (DiFonzo, 2005). Clearly the way in which the expert's opinion and its implications are communicated and explained plays an important role and has a significant impact on the jurors' perception of that piece of evidence (McQuiston-Surrett and Saks, 2007). Moreover, often the issue is that jurors may confuse opinions and facts and end up overweighting expert's testimony (DiFonzo and Stern, 2006). However, the novel aspect that these interviews unearthed is that high reliance on expert's accounts is shown also by judges, as their words described reliance on the experts' opinions, which are not questioned whatsoever if expressed with certainty. Understandably, it would be hardly possible for someone who lacks expertise in a given area to argue against an expert – the lack of knowledge would cause an obvious lack of credibility to those arguments. That does not necessarily mean though that experts' opinions

cannot be challenged (even by non-experts), yet this option does not seem to be taken into consideration. Nevertheless, these results are not surprising; they only provide further support to the claims, widely made by the previous literature, concerning the influence that experts' opinion may have on other individuals and on criminal cases' deliberations.

Unlike the first two judges, Judge 3 was not as explicit in his reference to forensic experts' testimony and he rather implied reference to them when referring to the technicality of the trial and to the necessity of showing the validity and the probative value of the evidence.

Judge 3: *'Well, clearly we move within a kind of trial that is technical anyway, besides the fact that it is murder, so either the defendant killed the victim or did not... But, to claim that that evidence leads to say that the defendant did kill the victim, we must explain why we accept them as valid and probative.'*

To clarify, this interview extract refers (in its original context) to the need for an expert's opinion in cases where the evidence that requires proof of validity and probative value is forensic evidence. However, this quote deserved mention in particular for its emphasis on an even more relevant aspect; the need to show a nexus between the evidence presented at trial and the final decision. In that lies the very nature of the principle of motivated verdicts and the words of Judge 3 show how deep-rooted that principle is in the minds of judges who have worked extensively in a criminal justice system that imposes that rule. In fact that principle results from a logical process: the evidence has to be evaluated because the decision must be based on it, therefore the jurors have to understand the evidence correctly and it is crucial that it is not misunderstood. All this is required by the fact that, in the words of the interviewee, they (the Italian jury) 'must explain why'. Thus the need to provide a written motivation, which is something that can only be done at the end of the deliberation, becomes a factor that affects the deliberation from the outset, permeating the whole discussion and the entire evidence assessment process.

3.3.2 Motivation

For the sake of terminological clarity, the term “motivation” is used, for the purpose of the present analysis, to indicate both the written document (integral part of the verdict, called “*motivazione*”) and the actual reason/s underlying jurors’ choices. The following discussion will first focus on the motivation in its first mentioned meaning, while later analysis (section 3.3.2.3) will refer to the motivation as a reason behind choices.

The need for Italian juries to motivate their verdicts is another broad topic that is of relevance to the present research investigation, as it represents another significant difference between the Italian and the British system. According to the Italian justice system, any judicial decision, taken by any magistrate or judging panel has to be accompanied by a (written) motivation (*Costituzione Italiana, art. 111, para. 1*). In Italy, thus, this is not a peculiarity of the jury trial introduced to confer greater guarantees to a judicial system that involves non-professionals. It is rather a general, fundamental rule that applies within each Italian court and to which all the judges as well as the other subjects who work within the justice system in Italy are widely accustomed. Given the absence of such principle in the British system, in analysing the differences, this was also a crucial topic addressed during the judges’ interviews. The intention in this case was to try and understand the various aspects of the rule to provide jury verdicts’ motivation as seen from the perspective of the judges; including what they think about the principle behind this need, about the duty of producing this written document, about the fact that the motivation is not required in other countries, and so forth. The discussion with the judges has indeed yielded interesting and at times surprising results. The themes which emerged while addressing this topic during the interviews were the following:

- Judges attitude towards the need to motivate verdicts;
- How do judges manage to write a motivation for a verdict on which they do not agree?
- When is the motivation (reason for choices) asked?

3.3.2.1 Judges' attitude towards the need to motivate verdicts

The need to motivate judicial decision is deep-rooted in the Italian legal system, so that it is frequently taken for granted by those people who work within it, to the extent that for them it becomes at times difficult to realise how there could be realities wherein that need is not recognised.

Judge 3 (referring to the British jury trial): *'Our reasons for appeals etc. are all grounded into the motivation... I really don't know how they proceed without motivated verdicts.'*

The explanation for this seems to have both substantial and formal/procedural grounds. Firstly, the written motivation is the "place" where the demonstration of a nexus between the evidence and the decision resides, and secondly, a "vitiata" (from the Italian definition "motivazione viziata", meaning "faulty/flawed motivation") motivation is the basis for an appeal, in fact it is the sole basis on which an appeal can be filed. However, as the following quotes will show, writing a verdict motivation is a very demanding job, which – despite being recognised as necessary – is not surprisingly seen as an unpleasant burden. It is bearing in mind these considerations and, thus, considering the judicial framework within which the Italian judges work, that their answers on this topic may be better understood. The first aspect that emerged was the different attitude that the interviewees showed towards this duty; the ones described below by the words of the judges are three very different attitudes, which show three different ways to perceive the motivation itself as a principle as well as the duty to write it:

Judge 1: *'It is obviously my job to write the verdict (Italian verdicts contain the motivation). It couldn't be otherwise after all, given that the jurors have no technical/legal competence to write that sort of document.'*

Judge 1's perspective appears fairly neutral and sensible in recognising that, given the need to provide a motivation, there is no alternative to the fact that the professionals on the panel are those who have to write it, as they are the only ones who have the necessary legal competence. This perspective does not raise particular issues or elements for discussion. However, judge 2 offered a different perspective:

Judge 2 (referring to the British jury trial): *'Their great advantage is that they don't have to motivate, so, once you are convinced, then you only respond to your own conscience [...] The advantage there is that you basically, yeah, you evaluate the evidence and then you get your idea, but you don't have to explain [...] I mean, if I find statements that are contrasting... contradictory... in the British system, say, once I'm convinced that one tells the truth and the other one lies, then I don't have to go on to explain why I believed one and not the other, which is what instead we need to do.'*

This answer is thought-provoking. Qualifying the absence of the requirement of a verdict motivation as an 'advantage' is quite a strong statement, to which various meanings may be ascribed. Similarly, referring to one's own 'conscience' as the only element to take into consideration when making the decision might appear an oversimplification of a situation that in reality should involve many more decision-making factors. Also, and perhaps most importantly, the last passage of Judge's 2 quote reported above deserves attention, in that it seems to move somewhat against the principle behind the need for motivated verdicts; that is individual belief is not enough to reach a decision on another person's fate, an objective link existent between the evidence and the decision has to be shown. Therefore, if judges come to the conclusion that not having to provide a motivation is an advantage, doubts arise on whether they actually perceive its importance as a tool to reduce the arbitrary power of unmotivated decisions and the increased risk and danger that those bring with them. Without neglecting this possibility, however, Judge 2's statements have, most likely, to be considered as the result of the weight that the unpleasant burden of motivation writing generally represents for judges. The following ironic remark confirms the negative connotation that writing a motivation has for judges.

Judge 3: *'It's our job. Surely they (the jurors) wouldn't be able to write a verdict... I sometimes threaten them (smiling/chuckling): "I'll make you write the motivation!"'*

Finally, to conclude on this theme, the perspective proposed by Judge 3 constitutes an interesting union of elements of the other two judges' opinions on the matter. It reasserts, on the one hand, the sensible consideration of the fact that judges are the only ones who would be able to write a legal document, which might also be seen as

the fundamental reason for their presence on the panel. It also, on the other hand, confirms the negative connotation that the redaction of this – at times overly lengthy – legal document has in the eyes of those who eventually will have to write it.

3.3.2.2 How do judges manage to write a motivation for a verdict on which they do not agree?

As it is clearly perceivable by the words of the interviewees, the negative connotation that the idea of writing a motivation report has is undoubtedly connected with the effort that creating such a long and detailed document requires. However, part of this negative connotation might have at times further causes. It is not difficult to imagine that the burden increases if judges find themselves required to write a motivation for a decision with which they do not agree. According to the Italian Code of Criminal Procedure, (CPP, art. 527, para. 3) the verdict decision has to be taken by the majority, and given the composition of the Italian jury (six lay members and two judges), it may well happen that the majority is constituted by the lay jurors only. In other words, if the jurors come to a conclusion that does not reflect the judges' belief, the latter have no power to change it, yet they still have to write a motivation for the verdict.

Judge 2: *'That's another big difference with the British system, because I can find myself in a trial, say, in minority regarding the decision, so when I have to write its motivation, I have to motivate something I'm not convinced of and that is very hard [...] and then the motivation is tailored on the final decision, it must be and if it is somehow inadequate, then this will come up in appeal, obviously.'*

Judge 2, maintaining his position of preference toward certain aspects of the British system, highlighted the paradoxical nature of providing a motivation for a decision that goes against one's belief. Indeed, writing a motivation means providing an account of the logical, explanation-based reasoning that, on the basis of the facts and evidence of the trial, has led to the final decision. Although it is the panel's decision, it is only the judge who has to provide an account of the motivation and describe the reasoning. However, when people engage in reasoning they normally undergo a process of selection of ideas and beliefs that leads them toward their decision, so that their arguments and conclusions are eventually mutually reinforcing (Mercier and Sperber,

2011). Consequently, it is easy to understand how difficult it must be to propose arguments that support a decision that a person would not have made themselves. Those arguments, in fact, would be the ones that the decision-maker would have discarded, in order to come to a different conclusion.

Furthermore, Judge 3 offered a view that takes into account an alternative outcome:

Judge 3: *'It never happened to me personally... But I know that it happened to colleagues; they didn't agree and had to accept a decision different from the one they would have made... In some cases there are the so-called "suicidal decisions" (sentenze suicide), so the judge will write a motivation that they know won't work in the further degrees, in appeal.'*

Judges might certainly make use of this loophole to ensure that jurors' decisions on which they do not agree are short-lived. The mechanism is straightforward; the appeal in the Italian justice system is grounded on "vices" (flaws) of the motivation, therefore creating a faulty motivation is the most "effective" way to see a verdict appealed. Although Judge 3 was referring to colleagues' experiences and only mentioned the "suicidal decisions" describing by hearsay the use of this tactic, this reality is unfortunately not completely unheard of in the Italian judicial system. A judicial sentence is defined as 'suicidal' when it shows incompatibility between the motivation and the decision, which happens in cases of so-called clear illogicality of the motivation (Tonini, 2011). Clearly, the possibility of using this instrument is neither recognised nor legitimated formally. It also implies, on the judges part, that their opinion is correct and necessarily more reliable than the jurors', which partially contradicts some of the previous statements around jurors being 'sensible' and 'just like them'.

To conclude on this theme, it is worth noting that the interviewees were also asked how they carry out this process for the production of the motivation; whether they take notes, how they remember the information, the various passages of the reasoning that they need to include in the written document. In response they argued:

Judge 1: *'No, I don't take notes... The motivation naturally comes out from the discussion just because each matter... each opinion are motivated. I don't need to take notes... It's a logical reasoning.'*

Judge 3: *'No need to take notes, I remember [...] many times they say sensible things, so... I mean... we can easily report their thoughts. Not the individual's thought (that is not allowed and it has to be kept secret) but, if in a given occasion someone has said something sensible that led us towards the decision, well, clearly that becomes integral part of the motivation.'*

The difficulty of the cognitive effort that judges are faced with in order to write a motivation for a decision with which they do not agree appears clear. Accordingly, given the extensively proven fallible nature of human memory and its tendency to let information be forgotten as well as leaving room for information to be “fabricated” (Ainsworth, 1998; Loftus and Loftus, 1980; Loftus, 1996b), doubts arise about the concrete possibility for judges to remember all the necessary elements to include in a motivation. When carefully considering these aspects, it cannot be excluded that judges' presumptions in regards to the “superiority” of their opinions (as emerged from the data) may tend to give them priority even if attempting to provide a comprehensive, accurate account of the group reasons rather than their own.

3.3.2.3 When is the motivation (reason for choices) asked?

As specified earlier, while the focus of the analysis conducted so far was the motivation to be intended as the written legal document that is integral to any Italian judicial decision, the discussion will now concern the motivation as a reason underlying choices. Indeed, considering that the written motivation should be the result of the reasoning conducted during the deliberation, it appeared relevant to ask the interviewed judges what, if any, rules are there for inquiring about those reasons during the discussion. Once again, the absence of set regulations about how to conduct the deliberation discussion was recognised, hence it was presumable that the judges' role would be discretionary and, therefore, it is important to investigate potential similarities and/or differences in their management/directive decisions.

Judge 1 confirmed that she applies the rule imposed by the Code of Criminal Procedure for the voting stages throughout the entire deliberation process, conferring a certain structure to the discussion. It is perhaps the very nature of this structured approach that brings with it the opportunity for her to consult the individual jurors about their opinions and motivations, quite constantly, in the various stages of the discussion.

Judge 1: *'The motivation for their choices is always asked to the jurors, every time they express an opinion and not only when they vote. That's obvious, after all... in any discussion... once you express an opinion, you have to motivate it. And this is important in the deliberation phase, so the members of the panel can understand the other members' choices and consider aspects that maybe they hadn't considered until that point.'*

From the words of Judge 2 an idea of greater flexibility came across rather clearly, as it seems that the (self-imposed) rule is in this case to consult the jurors at the end. Yet, as shown below, the judge carried on adding to his previous statement – with the tone of someone who is stating the obvious – a consideration on the natural instinct that leads to investigate the reasons of someone's choice.

Judge 2: *'The individuals are consulted at the voting stage... Well yeah, once a discussion starts, it's normal that one asks for what reason...'*

Similarly, Judge 3 reported his choice to consult the jurors about their motivations at the end and he also emphasised once again the existence of a formal attempt to not influence the jurors by letting them speak first as well as the clarity with which both opinions and motivations are expressly required from the jurors.

Judge 3: *'Basically that happens when we are about to decide, because that's when they are involved in the most incisive way in the trial and at that point I say... before I start speaking or my colleague does, I say: "So, each of you now please express their opinions and tell me why".'*

Broadly speaking, a fact that emerged from all the judges' answers in this instance, and indeed in several other occasions during the interviews, was their perception of some elements of the decisional process as obvious (e.g. the fact that one would naturally

ask about the reasons underlying a choice or, similarly, that verdicts need to be motivated). However, as demonstrated by the existence of systems that work under different rules, that “obviousness” is not an objectively recognisable element (i.e. British jurors/judges might think differently). This leads to the conclusion that the interviewees’ reactions might find their foundations in the development of a personal/professional background built inside a given system, which eventually has shaped the *forma mentis* of those people who were professionally trained within it by distortion of the reasoning process, so called “professional deformation” (Langerock, 1915). However, another sensible hypothesis could be that the attitude resulted from the fact that making choices on the basis of reasons is simply inherent in human behaviour, as a natural process which any decision-makers undergo (Shafir, Simonson, and Tversky, 1993). Therefore, while one can only speculate on the specific reasons that led the interviewed judges to perceive some aspects as obvious and somehow universally recognised, nonetheless a clear, logical conclusion may be drawn; if there is a motivation (even when unexpressed) behind anyone’s choice, that implies that any reasoning is in itself an explanation-based reasoning and, therefore, even the jurors who do not have to express their motivations indeed have some. The issue with the jury trial is that those motivations should be the “right” ones, i.e. those that, in a specific case, are relevant and legally acceptable.

3.3.3 Judges’ perception of their role

Following the analysis of all that emerged in terms of the judge’s role on the Italian jury panel and of the various aspects of the motivation required for jury verdicts, the analysis will now evaluate whether there are discrepancies between the actual role exerted by the judges and their perception of their role. Since it appeared clear throughout the interviews that the judges did not hold a full awareness of some aspects of their role, a certain degree of mismatch between their perception of it and reality was expected. The area wherein the majority of the discrepancies concentrated regarded dynamics of influence, which will be accordingly addressed in the following section.

3.3.3.1 *Influence*

Questions were asked with the intent to understand whether, during Italian jury deliberations, dynamics of influence are experienced within the panel, what could trigger them, whether they are regulated or arbitrarily enacted and how they are perceived by the judges. The starting point to reflect on this matter has to be the regulation of the Italian Code of Criminal Procedure, which, despite not setting strict rules on jury deliberation as a whole, appears to take a rather clear position on this particular issue.

The last stage of the deliberation process is the vote. This is also the only phase that is regulated by precise rules of the Code, according to which the votes are expressed by all members, starting from the youngest, and the judges are the last to vote (see art. 527 CPP regulation more extensively reported above on p.87). The initial question on the topic regarded this set rule and the judges all claimed during the interviews that they abide by this disposition and that they proceed with the vote as established by the law. However, it still seemed worth reflecting on the voting dynamics, as they offer an interesting source of information about how potential influence is disciplined and perceived through the lens of the legal system, and on the potential mechanisms of influence that can occur within the panel. On this point, unlike the matter of the voting procedure itself, the judges' responses acquire significant meaning and require cautious analysis.

The law, even if not expressly declaring so, establishes the above-mentioned voting procedure with the clearly perceivable intention of avoiding judges exerting influence on the jurors – 'lay jurors vote first' (CPP, art. 527, par. 2). This in turn implies that, even outside the express regulations, the law attempts to protect against the occurrence of dynamics of influence during the deliberation. And this intention is extended further – 'the presiding judge collects the votes starting from the youngest judge' (CPP, art. 527, par. 2). Both these regulations show the intention of the law with regard to mechanisms of influence and indeed both deserve discussion. To begin with, if the idea is that the judges are in a position that makes them particularly likely to be able, or even perhaps prone, to influence (consciously or not) the lay jurors, the same

cannot be thought with equal certainty about older and younger people. It is not clear on what basis it is believed that an older person is more likely to influence a younger one on a jury panel, therefore this rule appears merely and weakly grounded on the common rule of thumb whereby aging brings with it experience and wisdom.

Regarding instead the potential influence that the judges might exert on the jurors, the regulation appears to have more relevant reasons to exist. The possibility that the lay members of the jury are influenced by the professionals should be avoided, because, if that happens, the democratic principle that grants the popular sovereignty is contradicted. Accordingly, the rule established for the voting procedure should probably be extensively interpreted as going beyond the voting moment and applying to the entire deliberation process, in order to prevent dynamics of influence from occurring. In this respect the responses gathered through the interviews brought to light interesting elements. Attempting to discover what actually happens in the deliberation room and to identify differences between the actual and the perceived role of the judges, questions were asked in order to find out whether the interviewees believe that they influence the jurors or whether they believe influencing them is the right thing to do given their lack of knowledge/expertise in legal matters.

Judge 1: *'I don't think I adopt a behaviour that tends to influence the jurors. Following the rules of the Code, my vote is expressed at the end. During the discussion as well, the various opinions are expressed following that order; from the youngest to the oldest. This system not only protects the laypeople from the influence that we – professionals – might, even involuntarily, exert, but it also creates a discussion where the older/more experts don't prevail on the younger people... a discussion where, through the exchange of opinions, members of the panel consider also aspects that didn't take into account at first.'*

Judge 1: *'I can't exclude categorically that certain dynamics of influence are activated anyway during the discussion, but can definitely say that we professionals do try to limit their occurrence. However, even if it happens, I believe it's normal in social interaction, it'd happen in any similar context, that is people deciding together – also professionals.'*

Judge 2: *'It's obvious that – if the judges were to vote first, it's true, they could influence... they close the circle instead.'*

And when asked whether the same principle was followed also during the discussion:

Judge 2: *'Well, no... principally for the vote.'*

Judge 3: *'Leaving them free to express their opinion is the first thing. Then, in the end, if we don't agree, we try somehow to influen... to make them understand why, in our opinion, their reasoning has some weakness, some flaw.'*

Given the impulsive use of the word 'influence', then retracted, the following question was: *'So do you think that this is part of your directive role? They should be influenced then.'*

Judge 3: *'I wouldn't use that word, "influence"... I'd say, make them understand why, on the basis of our judicial system, that opinion they have is not correct. Influence them, no, 'cause I don't... it'd seem that there is a role of prominence... There is a role of direction actually.'*

It appears clear that the discussion about this theme unearthed very interesting truths. Despite the fact that the Code fundamentally prescribes the assumption of behaviours that, in general, should not let mechanisms of influence enter the deliberation process, the judges maintain their freedom. Indeed, while they all recognised (at least formally) that they aim to avoid influencing the jurors, it is clear from the differences in their answers that they discretionally choose how to try and achieve that aim; Judge 1 follows the vote procedure also during the discussion, whilst Judge 2 does not – neither of these two behaviours can be criticised, as there are no clear rules to this regard and therefore they are both formally legitimate. Yet, they give rise to different situations, which both seem to present advantages and disadvantages. It could be argued that in Judge 2 deliberations there will be much greater likelihood that the professionals' opinion – expressed at any time during the discussion, and thus potentially even before jurors' at times – will turn out to be influential and affect the lay members. By contrast, in Judge 1's case there appears to be less room for influence, yet it should also be considered that a discussion which follows the described fixed structure (with members' opinions expressed in a certain order) is very likely to lose a great deal of its natural flow and with that, possibly, still affect individuals' spontaneity in sharing their ideas.

Moreover, the answers provided by Judge 3 are relevant as well, as they offer a further perspective, which starting from the theme under analysis might then be extended to the judges behaviour as a whole. His first use of the word 'influence' (which he did not pronounce entirely, managing to stop himself) might have several alternative meanings. It could mean that he actually believes that the judges should in fact influence the jurors, but he thought it would not be appropriate to say so, perhaps because he is aware that this would somehow be against the formal rules. It could alternatively mean that he does not believe that influencing the jurors would be against any rules (there are actually no express regulations), but still consider inappropriate to admit that he finds it fair to influence them. Ultimately, it could also merely mean that he made a mistake using that word – he very clearly states afterwards that he would not use that word (even though he actually did).

When using interviews as a research method it is often very difficult to determine with certainty whether the interviewee is being honest (Leedy and Ormrod, 2013) and also in this case it would not be appropriate to draw conclusion of this sort. However, what can be concluded with a good degree of certainty is that from some of the reactions to the questions posed it seems that the judges, despite an apparent confidence, most times answered as if they never really had to think about those issues. Sociological research has shown that human beings are often not very accurate in their observation, understanding and interpretation of their own and others' behaviours (Marshall, 1996). This could explain some discrepancies between the actual role of the judges and how, on the other hand, that role is perceived by these actors of the Italian criminal justice system. In turn the existence of these discrepancies seems best explained by the absence of formal regulations which leave the judges free to decide what it is that should be done.

Indeed one of the first truths that the interviews revealed is that the judges actually possess a high degree of freedom in determining how to behave. The study, on the one hand, emphasised once again the discretionary nature of their role, yet on the other hand, also highlighted the existence of several commonalities among the choices individually made by each judge on how to behave. Similar behaviours could normally

be the result of individuals' training or compliance to set rules. Yet, in this case, the existence of behavioural similarities among the judges suggests that there must be a, loosely defined, set of common self-imposed rules that all the judges tend to follow. Such conclusion appears consistent with the psychological literature. Human behaviours are generally controlled in significant measure by externally imposed rules or – perhaps more accurately – by the externally established consequences of their actions, should those go against the rules (Mischel, 2007). Nonetheless, it has been asserted that, even in the absence of those external limits, individuals are prone to auto-regulate themselves. Although in such cases the consequences/sanctions are objectively absent, the tendency to behave in a certain way still remains as a response to personal rules, given the negative emotions that the violation of the self-imposed norms may generate (Bicchieri, 2013).

3.4 Conclusion

In a context where entering the deliberation room is prohibited, the study presented in this chapter provided a highly valuable source of information on deliberation dynamics, as depicted by the professional perspectives of judges who work inside it (in Italy). Audio-recorded, semi-structured interviews were thematically analysed and shed light on the characteristics of these professionals' role. Even in the absence of formalised rules, judges were mostly consistent with one another in their behaviour, which suggests the existence of consistent self-imposed rules (Bicchieri, 2013). The occurrence of a number of contradictions throughout the interviews showed either lack of awareness or inaccurate perception of their role on the part of the judges (Marshall, 1996). However, it was possible to observe, overall, the highly discretionary nature in the judges' exercise of their functions. Some of the advantages of the presence of these professionals were clear, yet, given their influential power, doubts arise about the threat that this power may constitute against popular sovereignty. Overall, the benefits of these findings were two-fold: firstly, they offered an opportunity to reflect on the role of professionals on mixed juries as well as on their (perhaps too highly) discretionary and influential power; secondly, the results became a

methodological instrument to be used in Study 2 in order to instruct the mock judge used in the experiments on the basis of realistic information.

CHAPTER 4

MOCK JURY EXPERIMENTS:

Comparing British and Italian (mock) juries

*'We think, it is true, but we think so badly
that I often feel it would be better if we did not.'*

Bertrand Russell

4.1 Introduction

The previous chapter described the first study conducted for this thesis which, through interviews with Italian judges who work on jury panels, provided crucial information for the comparison, between British and Italian jury systems' functioning. In addition to offering a better understanding of the internal functioning of the Italian jury, the interviews provided important information about the role of Italian judges, which was then used to inform the design of the experiment described in this chapter. This study was designed and developed in order to understand and assess the potential impact of two variables on the deliberative process of jury decision-making – specifically the presence of judges and the requirement for motivated verdicts.

Manipulating the two variables so that mock juries could deliberate either according to the British jury trial rules (only lay people, rendering unmotivated verdicts), or according to the Italian jury trial rules (mixed jury, rendering motivated verdicts), the participating groups deliberated on a fictional trial in order to reach a unanimous verdict. The audio-/video-recording of the mock juries' deliberations along with participants' responses to pre-/post-deliberation questionnaires provided a considerable amount of both qualitative and quantitative data that are reported and discussed in this chapter.

4.2 Method

Through the employment of what previous literature has determined to be the most appropriate empirical method to investigate jury decision-making (Hastie, Penrod and Pennington, 1983; MacCoun, 1989; Ellison and Munro, 2010a), the present study extended the understanding of deliberation in the field of jury decision-making. Accordingly, while general methodological considerations on both studies and the interconnection between them was provided in Chapter 2, a detailed account of how the mock jury experiments were designed and conducted is provided here.

4.2.1 Participants

4.2.1.1 Nature of the sample and recruitment – Lay participants

The sample used in this study was a combination of students and members of the general public. Justifications for these participants' recruitment were provided in Chapter 2 (section 2.4.2), where general methodological choices are explained, and will be briefly recapped here. Experiments with British mock juries (panels that deliberated according to the British jury system procedures) were conducted in the UK, whilst experiments with Italian mock juries (panels that deliberated according to the Italian jury system procedures) were conducted in Italy. In both countries the study was advertised through the use of typical recruitment methods, such as social networks and word of mouth (Bankert and Amdur, 2006). In the UK, the experiment was also advertised through the distribution of flyers (Appendix D) and through brief oral presentations from the researcher during lectures. All these methods are considered to be some of the more commonly used types of direct advertising in research (Bankert and Amdur, 2006).

When the study was advertised by the researcher in person at the beginning of a lecture (in the UK), students were given a sign-up sheet, already organised in timeslots, so that they could select a convenient date and time to take part in the experiment. Willing participants who became aware of the study from the flyers were asked to email the researcher, who would then randomly allocate them to a few date/time slots in order to let them choose depending on their availability. A similar procedure was

used in Italy, where potential participants who became aware of the study through social networks and word-of-mouth advertising (Bankert and Amdur, 2006), were also offered alternative dates/times in order to meet their needs and availability. In all cases, to encourage participants' attendance, those who volunteered to participate were sent follow-up invitations and reminders (Krueger, 1994). As a result, in the British condition, mock juries were mostly comprised of jury eligible students (77%), given the researcher's connection with the university environment and the consequent higher accessibility of this type of participant. Jury eligible members of the general public constituted the majority of the mock jurors used in the Italian condition (79%), where the lack of connection with the university environment reduced the variety of effective recruitment methods. Notwithstanding the awareness that these differences in group composition are a limitation, given the difficulties in human participant recruitment and based on previous research findings that 'juror characteristics are rarely determinative of verdict outcome' (Ellison and Munro, 2010a, p.76), this was not considered to be an overly influential limitation. However, a potential impact of limitations related to the sample used in this study has not been disregarded and further considerations on this point are provided in Chapter 5 (section 5.7), where strengths and limitations of the entire research project are discussed.

4.2.1.2 Nature of the sample and recruitment – Professional participant

With specific reference to the Italian condition, further clarifications regarding the nature and recruitment of the sample need to be made. In order to recreate the dynamics and conditions of Italian jury deliberations, it was considered essential to take into account the mixed composition of Italian juries and attempt to replicate them in the experimental condition. Therefore, since the presence of professional judges on the Italian jury is one of the key factors of jury deliberations, this variable needed to be sufficiently recreated. Previous empirical endeavours have acknowledged the difficulties of recruiting such professionals and resolved it by modifying experimental conditions. For instance, Fujita and Hotta (2010) sacrificed external validity and generalizability of their results, deciding to neither ask professional judges to participate, nor ask participants about their occupation. Clearly,

this approach would not have been appropriate for addressing the research questions posed by this thesis. Thus, acknowledging the difficulties of involving professionals who are not typically engaged in the research process despite their unique insights/position (Temkin, 2000; Bleich and Pekkanen, 2013), for the purpose of this stage of the investigation, an alternative solution was adopted.

The connection of the researcher with the Italian legal/courts environment, due to her past working experience as a lawyer, made it possible to recruit a former colleague to play the role of a judge on the Italian mock juries. The legal knowledge that judges and lawyers have in the Italian criminal justice system is equal: professionals in both categories undertake the same university qualifications and are required to hold the same knowledge of legal matters, which they approach from different perspectives, according to their different roles. For this reason, real judges were needed (and could not have been substituted by any other professionals) for the first study, in that it was aimed at understanding aspects of their role that only they would know; whereas, for the second study, the competence and expertise of a lawyer were considered sufficient to the scope. Indeed, all that was needed for this study was, on the one hand, knowledge and expertise in legal matters and, on the other hand, awareness of the role of a judge when working on jury panels. While any lawyer already possesses the former, the information acquired from the first study (Chapter 3) enabled the researcher to equip the participant lawyer/mock judge with the latter.

Five lawyers were contacted, and two agreed to participate. However only one was selected for the experiments. This choice was made for methodological reasons, after careful considerations of benefits and limitations of having only one mock judge across all the panels. While it is true that, in reality, judges do not deliberate on the same criminal case multiple times, it is also true that for an experimental design this deviation from reality did not appear to be detrimental and was, in fact, beneficial. Indeed, as a general rule in research methods' literature, 'to make sure that unintended variables do not confound a study, many potential variables are kept constant, or the same, across conditions' (McBurney and White, 2009, p.122). In this study, the "identity of the judge" (with all that this concept contains and implies: e.g.,

attitude, competence, assertiveness, charisma, etc.) was a potential variable, which, if not held constant across the Italian mock juries, might have generated unintended effects. Therefore, the same lawyer was asked to play the role of a judge on all five Italian mock jury panels.

The lawyer who agreed to participate confirmed her availability to partake in all five mock jury deliberations and also agreed to attend a pre-emptive briefing session, during which she was instructed on how to undertake the role of a judge. Obviously, no legal information had to be provided regarding the judicial proceeding, given that, as a criminal lawyer, she attends trials on a daily basis. However, instructions regarding what happens inside the deliberation room were needed. These instructions, obtained in part from the previously analysed judges' interviews (Chapter 3), were summarised in order to provide the mock judge with enough information about how to conduct the discussion and how to manage the interaction with the lay participants.

4.2.1.3 *Sample Size*

The overall sample obtained for the mock jury experiments comes as a result of the endeavour to have ten mock jury panels (five in the British condition; five in the Italian condition), with six mock jurors on each panel. It was taken into account that a six-member mock jury panel does not exactly mirror the actual composition of real juries, neither British nor Italian, since such panels are made up of, respectively, twelve and eight members. However, as it has been argued, a certain degree of restriction in jury size is required to grant practicality of empirical studies (Ellison and Munro, 2009) and, therefore, the idea of smaller sized mock jury panels was contemplated and accepted for the purpose of the experiments. Moreover, in light of existing literature, this does not appear to be an overly problematic issue, since previous studies have found no (or only few) significant differences between six- and twelve-member mock juries (Hastie, Penrod and Pennington, 2002). In Hastie, Penrod and Pennington's (2002) review of studies on mock jury size, studies were found that registered no significant differences between six- and twelve-members mock juries (Davis et al., 1975; Buckhout et al., 1977 cited in Hastie, Penrod and Pennington, 2002). Yet, another study in the review recorded a difference in deliberation time, which was shorter in six-people juries.

Nevertheless, this did not affect the average contribution to the deliberation on the part of individual jurors, as this was found to be equivalent in mock juries of both sizes (Friedman and Shaver, 1975 cited in Hastie, Penrod and Pennington, 2002). At closer inspection, it may be considered that, even when differences were found, they seemed not to be greatly influential or even favourable for the purposes of this study. The latter study's findings, for instance, may indeed be seen as advantageous to the practicality of the experiments, especially when considering that mock jurors' deliberation times could never equate to real ones.

The attempt to have exactly six individuals per panel was successful in seven out of ten groups. This was due to uncontrollable circumstances, such as a few participants not attending. Research methods literature recognises this as being an expected difficulty in research using groups (Krueger, 1994) and, therefore, despite measures taken to avoid it, no-show-ups could not be completely avoided. Luckily, this never resulted in the need to cancel an entire mock jury session, because the few no-show-ups were considered acceptable, in that they consisted of:

- One participant missing in mock jury B2;
- One participant missing in mock jury I10;
- Two participants missing in mock jury B3;

Given the very small number of no-show-ups and the importance of voluntary participation in this sort of activity, all mock jury sessions still took place. Therefore, a total of 52 participants took part in the experiments; 27 in the British condition and 25 (24 lay mock jurors and one mock judge) in the Italian condition. Five more participants took part in a pilot study (described in section 4.2.4.1), but were not included in the final analysis.

4.2.2 Design

In order to investigate the potential effects of the two above-mentioned independent variables on jury deliberation dynamics and answer the second set of research questions (Figure 2.2), the recruited participants were divided in ten groups and

required to deliberate on a fictional trial scenario. Five groups deliberated according to the British system's rules: only lay participants took part in the deliberation and had to render an unmotivated verdict; five groups deliberated according to the Italian system's rules: lay participants and one professional participant playing the role of a judge deliberated together, having to reach a motivated verdict. This generated two experimental conditions, which are referred to as the "British condition" and "Italian condition". Accordingly, mock juries used in the experiments are identified by a code: B1, B2, B3, B4, B5 are British mock juries; I1 I2, I3, I4; I5 are Italian mock juries. Similarly for the jurors: J1, J2, J3, etc. Hence, to indicate a specific juror on a specific jury, a combination of codes is used: for example, J5 I3 will be juror number 5 on the Italian mock jury number 3.

As suggested by the methodological approach adopted by Ellison and Munro (2013), deliberations were audio-/video-recorded and pre-/post-deliberation questionnaires were given to participants for completion before and after participation. A qualitative methodology was utilised to identify themes within the mock jurors' deliberations and quantitative analyses were further carried out with data acquired mainly from the questionnaire responses.

4.2.3 Materials

Regarding the experimental stimulus, all the mock juries in the experiments were presented with the same crime case scenario. Previous published research has demonstrated that this is an effective approach, in that it illuminates how different groups, when faced with the same stimulus, may either reach the same or the opposite outcome, taking various routes (Ellison and Munro, 2010a). Additionally, in an effort to choose an experimental stimulus that could guarantee practicality and convenience, whilst also being sufficiently ambiguous, complex and engaging, the trial scenario for this study was created as follows.

Rather than designing a completely new mock trial stimulus, a pre-existing source was utilised as a starting point and modified according to the purposes of the study. The source was a crime scenario created by the University of Hull, funded by the Higher

Education Academy, and made available as an open source for teaching and research purposes in Forensic Science and Law (JISC Open Educational Resources Programme, 2011). The document was originally designed as an exercise for Law students to practice developing a defence strategy and was presented in the form of a crime case file, containing several elements related to a hypothetical murder/rape case. It includes the police case report, list of exhibits, evidence, testimony, suspects, and crime scene photos (see Appendix E for the original PDF file).

As anticipated, the material was not suitable in its original form and required adjustments to be made for the purposes of the experiments in the present study. First of all, due to the above-mentioned original aim of the exercise, it lacked a defence strategy, which instead needed to be designed. In order to grant validity to the defence strategy, a few steps were taken. Firstly, feedback from students was collected throughout the years, when the material was used, by the researcher and her supervisor, in its original form for teaching purposes. In such instances, while students developed their own defence strategies, notes were taken from the researcher of the most frequently mentioned pieces of evidence, as well as of the most commonly proposed arguments. This provided a preliminary broad indication of how a thought-provoking defence strategy should be designed. Starting from this basis and, additionally, making use of the researcher's professional background as well as of the advice of a lawyer, a defence strategy was designed and incorporated into the stimulus. Subsequently, to give the mock jurors the experience of a complete trial, prosecution and defence closing arguments were also scripted and added to the stimulus. The modified document was not given to mock jurors (as it would be to the Law students undertaking the original task). It was considered that, if given the document, mock jurors would have different access to the information and would, presumably, spend different amounts of time reading through it, with some participants being more cautious than others. Therefore, in order to maintain consistency in the information acquisition phase, mock jurors acquired knowledge of the facts of the case all at the same time and in the same way: that is, the case was shown and described to the mock juries through a Prezi presentation (Prezi.com – see Appendix F).

Furthermore, as noted above, mock jurors were required to also respond to questionnaires. An initial brief questionnaire was designed to collect demographic information (Appendix H). In addition, pre-/post-deliberation questionnaires (Appendix I) were designed to collect mock jurors' individual views about various aspects of the deliberations and, among other things, to offer them the opportunity to provide their individual opinions before and after the discussion. It was considered that this would offer the researcher the valuable opportunity to obtain a deeper understanding of individual verdict preferences, which could otherwise easily be lost during the deliberations. In such contexts, more silent members' preferences would have run the risk of remaining hidden behind the leaders' louder voices, converging into verdicts, only apparently unanimous.

4.2.4 Procedure

Once the trial scenario was adapted for the purpose of the study and the questionnaires were developed, the experiment was piloted with a group of five Criminology distance learning students attending a residential study school at the University of Leicester. The need for a pilot was strongly felt since multiple elements and the functioning of them altogether had to be tested. It was necessary to ensure that the hypothetical case was engaging in the way it was presented, and ambiguous enough to prompt a discussion that could offer elements for analysis. In addition, the clarity and effectiveness of the questionnaires required assessment. Accordingly, the feedback received from the pilot participants was used to improve aspects of the experimental material and to evaluate the complexity and controversial nature of the case. As a result, the wording of some questions was slightly modified to make them clearly understandable and unambiguous, and a few marginal aspects in the description of the case were amended – for example, explicit reference to the fact that the ballistics examination results were inconclusive was added, as the absence of this detail seemed to generate confusion in the pilot. The hypothetical case scenario presented to the pilot participants prompted heated discussion, arising from different initial verdict preferences and was, consequently, considered appropriate for the actual experiments.

The following procedure was used for both the participants in the pilot study and, following necessary amendments, the actual mock juries in the British condition. Slight changes occurred in the Italian condition due to the presence of a mock judge (which was not specifically tested in the pilot), although, as will be shown, they related mostly to the deliberation phase itself and not the presentation of the case.

4.2.4.1 Pilot and Mock Juries in the British condition

Participants were invited to report to the Criminology Department Annex at the date and time they chose. Before starting, they were provided with a recap of all the general information needed along with some more specific instructions regarding the different steps of the activity (description of the case, pre-deliberation questionnaire, discussion, post-deliberation questionnaire, etc.). Each of the mock jurors found at their respective places at the table an informed consent form (Appendix C) and the first brief questionnaire, through which they provided demographic information about themselves (Appendix H). They were also told that on the back of the first questionnaire sheet there was a box for their notes about the case (as indicated by the Crown Court Compendium, 2017), which they could use to write down relevant aspects of the case during the presentation.

After mock jurors' completion of the first questionnaire, the presentation of the case began. Participants were warned that they could ask the researcher any questions about the experimental activity, but they could not receive extra information about the case. This way, for the sake of consistency, it was ascertained that all juries deliberated with exactly the same information and any difference in the stimuli could be excluded. The case was presented by the researcher, using a Prezi presentation, so that participants were exposed to a number of stimuli: spoken and written words, photos, etc. (Appendix F). In order to further ensure consistency of stimuli across groups, the description followed a script that was used by the researcher (although not given to the jurors) to ensure all the mock juries were provided with the exact same amount of detail. The presentation of the case closed with the reading of prosecution and defence closing arguments, which provided the respective evidence-based reconstructions of the facts of the case (Appendix G).

Following the presentation of the case, mock jurors were given a second questionnaire. This pre-deliberation questionnaire (Appendix I) was designed mainly to obtain individual jurors' verdict preferences, before the discussion started, in order to be able to verify whether and how these preferences might be swayed by the deliberation dynamics. The pre-deliberation questionnaires were collected after completion and before the beginning of the deliberation.

Mock juries were then given a few instructions. They were told that they could freely decide how to conduct the deliberation, that is no rules were imposed on whether they should elect a foreperson or whether they had to take votes before the discussion started. This deviation from real world settings was purposely planned in order to observe potential trends in groups tendencies in terms of leadership dynamics and approach to deliberations. Since these two factors were of particular interest for this research purposes, especially considering their hypothesised impact on deliberation dynamics, this methodological choice was made with the awareness of paying the price of a slightly decreased realism in order to prioritise specific needs of the investigation. Therefore, the only express requirements were for the juries to decide on both counts (murder and rape) and to aim to reach a unanimous verdict. With these last instructions and after having ascertained that everything was clear and there were no questions, the researcher left the ("deliberation") room with the agreement that, once the verdict had been reached, one of the members of the mock jury contacted the researcher who would then proceed to distribute the last questionnaire.

The post-deliberation questionnaires, as indicated above, obtained information about jurors' perceptions and opinions about deliberation dynamics and produced elements of analysis on whether and how individual verdict preferences of mock jurors changed during the discussion. The questionnaires, along with the video-recordings of the deliberations, gave a comprehensive view of the nature and potential reasons of these changes of mind in mock jurors and provided a variety of data which, through triangulation, were used to answer the research questions. To conclude, once the post-deliberation questionnaires were completed and collected, the mock jurors were

asked to confirm their email addresses where they would receive their Amazon Voucher (see Ethics section below about incentives) and the sessions ended.

4.2.4.2 *Mock Juries in the Italian condition*

As noted, the experimental procedure changed slightly for mock juries in the Italian condition. The presence of a mock judge and the requirement of a motivated verdict necessitated some adjustments. Mock jurors gathered in the “deliberation” room were made aware at the outset of the presence of a professional participant; they were told that the professional was a lawyer who was going to play the role of a judge, given her expertise and knowledge of legal matters. It was also clarified what that role consisted of, so that the lay members knew what to expect from the mock judge during the course of the deliberation. After this brief clarification, the same procedure as the British condition was followed, with the completion on the part of mock jurors of the first questionnaire, the presentation of the case and the completion of the pre-deliberation questionnaire. Different to the British mock jurors, those in the Italian condition were instructed that they had to motivate their verdict.

The way in which the deliberation was conducted also slightly changed. The procedure was adapted in line with the findings from the judges’ interviews (Chapter 3), so that the mock judge was instructed to act as real judges do. This specifically meant that a written summary report was provided to the mock judge, who went through it at the start of each deliberation and kept it there in case there was need to double-check information during deliberation. Legal principles were explained, motivations were expressly asked of the lay jurors (the mock judge asked everyone what their reasons were at the outset and throughout deliberations). Moreover, the post-deliberation questionnaire in the Italian condition, according to the needs of the analysis, was slightly different, in order to be adapted to the presence of a judge, and additionally to require jurors to provide a motivation (Appendix I). These formal/procedural aspects that changed in the Italian condition produced, as expected, concrete differences. However, since these differences were indeed at the heart of the present investigation, they will be addressed in detail in the Analysis section of this chapter.

The only other difference between the two conditions that deserves mention is the language spoken. The English language was used in the British condition, and the Italian language was used in the Italian condition. This included both the stimuli (created in English, and then translated in Italian) as well as the language spoken by the researcher and the participants. To dispel any doubts or concerns in terms of issues of consistency of the analysis due to the language difference, similar considerations to those presented regarding the judges' interviews in Chapter 3 (section 3.2) had to be made. In order for the mock jurors to act as realistically and naturally as possible, it was considered much better to expose them to stimuli in a language they felt familiar with and to allow them to speak their first language, without posing further unnecessary obstacles to the realism of the experiments.

4.2.5 Ethics

To ensure compliance of the present study with ethical standards, a research ethics application was submitted to, and approved by, the University of Leicester Ethics Committee. At the recruitment stage, only general information about the study was given. Potential participants were given indications about the nature of the activity they would be required to carry out and the duration of participation. They were also offered an incentive to participate in the form of a £5 Amazon Voucher. Consistent with literature that considers this as a legitimate approach, which offers the opportunity to "thank" people for their time (Grady, 2005; Cleary, Walter and Matheson, 2008), Amazon vouchers were chosen as a popular form of incentive in research studies, mainly because they are 'easy to administer [...] easy to redeem and of widespread appeal' (Poynter, 2001, p.14). Indeed, participants in the present study could very easily redeem their vouchers online through a link that would be sent to their email address following their participation.

Willing participants were required to sign an informed consent form (Appendix C) which further clarified the nature of the activities they were required to undertake. Although participants were reassured that they would experience no harm as a consequence of their participation in the study, they were also warned of the presence

of material of sensitive nature (e.g. fictional crime scene pictures). Participants were informed of their right to withdraw at any time and were provided with the researcher contact information in case they had questions, doubts or any types of enquires. Participants were also ensured confidentiality; their names do not appear anywhere in the thesis and each participant was given an anonymised identification code. Furthermore they were assured that audio-/video-recording would be listened/watched by the researcher only and would be destroyed after completion of the research in order to protect participants' identities.

4.3 Analysis

Consistent with the aims of the study and with the nature of the data that it enabled to be collected, the analysis conducted had both qualitative and quantitative elements. General considerations on reasons, benefits and limitations of the employment of a mixed methods design have been already proposed (Chapter 2) and therefore this section, to avoid repetition, will focus on how the analysis was conducted. In particular, considering that there were two main sources of data (videos and questionnaires), qualitative content analysis was carried out with data gathered mostly from the videos, while descriptive and inferential statistics were performed on numeric data emerging mostly from the questionnaire responses. The analysis of both resulted in a triangulation design, employed, as suggested by the research methods literature, to either compare and contrast or to corroborate quantitative statistical results with qualitative findings and *vice versa* (Creswell and Plano Clark, 2007).

4.3.1 Stage 1: Identification of themes and coding process

According to research methods literature, for the purposes of qualitative content analyses, an initial categorisation of data into themes can be obtained "*a priori*". As Bernard and Ryan (2010) explained, themes normally derive both from the data gathered and from a prior understanding of the phenomenon under analysis. The former include all those themes that emerge from the data and that is impossible to anticipate; the latter are, instead, deduced *a priori* either because they arise from characteristics that are integral to the phenomenon studied, or because they come

from well-established conclusions reached by the literature on the topic, or else because they are directly connected to the research questions and to what exactly the researchers are attempting to understand. Following this approach, aimed at deductively and inductively identifying themes, the coding procedure was designed. The coding process was conducted through the employment of the template analysis technique (King, 2012).

King (2012) emphasised that the template analysis technique, widely used in the context of different methodologies due to its flexibility, presents the benefit of conferring a high degree of structure to the analysis. This was considered particularly suitable for the present investigation, given the high volume and multifaceted nature of the data gathered. The adoption of the technique involved the development of a coding template, namely a list of codes representing themes potentially identifiable in the data (McDonald, Daniels and Harris, 2004). According to what was anticipated and to what frequently occurs with the use of this technique, initial coding was broadly constructed on the basis of themes that were strongly expected to emerge (Brooks et al., 2015). However, following previous research recommendations (McDonald, Daniels and Harris, 2004), where necessary, *a priori* codes and themes were amended and new ones (*a posteriori* codes) were developed throughout the analysis, as familiarity with the data increased and new relevant codes and themes emerged. During the following stages of the analysis, as King (2012) suggests, notes were taken of all the elements that appeared to be meaningful in relation to the research questions. When those elements fell within one of the *a priori* codes, they were inserted in the related section and contributed to the development of a broad theme. When they did not, they were either discarded (if, for example, they appeared only once and there was no possibility of establishing any connections or finding any patterns across groups regarding that specific element), or inserted at the top of a new coding section. This resulted in a hierarchical (with higher and lower-order themes) template used to code/organise the qualitative data (see the coding template in Table 4.1).

4.3.2 Stage 2: Questionnaires' responses – preliminary exploration

For practical reasons, since it was the researcher's intention to transfer the paper questionnaires into electronic format for greater clarity and faster, more effective consultation, the analysis started from that source of data to then move on to the video deliberations. Since it has been widely recognised that the activity of transcribing has to be considered an integral research activity in its own right rather than a technical aspect that precedes the analysis (McLellan, MacQueen and Neidig, 2003), in this study the activity of transferring the hardcopy questionnaires into electronic format was considered an important step in the overall analysis. Accordingly, the purpose of this activity was producing an easily consultable backup copy of the data, while becoming familiar with the data, reflecting on the questionnaires' responses as well as taking notes of some of the most evident trends and patterns that appeared to be emerging across groups. This first contact with the data became one of the bases for further confirmation of the validity of the coding system for qualitative content analysis as well as for the creation of variables for statistical analysis.

4.3.3 Stage 3: Video deliberations – first view

Subsequently, once the transcription of the questionnaires was completed, the analysis shifted to the other source of information; the video deliberations. As well as for the questionnaires, the intention underlying the first approach to the videos was to begin familiarising with the data: understanding how the discussions went, identifying the most evident similarities and differences across groups overall, and between groups in the two different conditions, etc. All the videos were, therefore, watched a first time and notes were taken sparingly, as quick reminders of aspects that deserved attention in relation to the research questions (King, 2012) and would be confirmed or discarded throughout the following stages of the analysis.

During this preliminary view, precise information on actual length of deliberations (section 4.4.2.2), number of participants per group (no-show-ups had to be detracted from the expected participants' number – section 4.2.1.3), and individual jurors' initial and final votes were recorded. This final piece of information was, in fact, already

gathered while transferring into electronic format the questionnaires' responses (Appendix I, Q. 1, pre-deliberation questionnaire; Q. 3, post-deliberation questionnaire); the extra step of recording it from the videos was taken to then compare and double-check whether mock jurors' questionnaire responses corresponded to their claims during the deliberations. Once again, this approach, which was adopted throughout the various stages of the analysis, enabled improved validity of the findings and further demonstrated the effectiveness of the use of pre-/post-deliberation questionnaires. They indeed proved to be, on the one hand, an effective tool to verify information provided by the participants and, on the other hand, a way to encourage the mock jurors to voice their opinions, since they knew they had already declared them in writing.

4.3.4 Stage 4: Video deliberations – second view

Subsequently, all the videos were watched a second time. The second view differed significantly from the first in terms of approach, that is, time, caution and precision in analysing the data. The aim of the second view was to conduct a meticulous observation of the deliberation dynamics that the videos had been able to capture. In this context, it was possible to identify commonalities in the development of the discussions, group dynamics, leadership, private/public conformity, etc., which in turn laid the foundations for the identification of occurrence of phenomena, emergence of recurring themes, and thus for the qualitative aspect of the analysis. With this aim, during the second view, deliberations were transcribed for later analysis. Differently from other studies, however, and in accordance with the purpose and specific characteristics of this empirical endeavour, video deliberations were partially, or in the words of linguists, "broadly" transcribed (Gee, 2011). This choice was made in an effort to obtain the most effective balance between data required for the analysis and time constraints, as well as in light of what previous literature suggests in this regard.

Transcription is indeed considered to fall within two different categories: full or partial (Wasamba, 2015). It has been compellingly argued that 'researchers must decide whether their analysis is best supported by transcription or by researchers' notes

derived from or supplemented by a review of the tapes' (Patton 2002:380). According to Gee (2011), this is due to the fact that such analyses are not generally based on all the spoken elements present in a verbal exchange, but should rather focus on elements that are meaningful and relevant to the specific analysis conducted. These judgments of relevance – what should/should not be included in a transcript – are referred to the analyst/researcher's view of how interactions unfold in the specific context under analysis. Moreover, the selection of text for transcription should be carried out on the basis of a consideration of the actual contribution that it will make to the overall analysis (Strauss and Corbin, 1990) or, in other words, 'the level of transcription should complement the level of the analysis' (Drisko, 1997, p.190).

As a consequence of these considerations, 'any speech data can be transcribed in more or less detailed way [...] ultimately it is the purposes of the analyst that determine how narrow or broad the transcript must be' (Gee, 2011, p.117). Indeed, the purposes of the present analysis did not appear to require a full/narrow transcription of the video/audio-recorded deliberations. First of all, the present study did not intend to focus on speech content itself, because they would not have directly responded to the needs of this research project; secondly, the video deliberations were only one source of information in this study and were intended to be used in conjunction with the questionnaires.

The chosen use of partial transcripts appears to be further supported by McLellan, MacQueen and Neidig (2003) who claimed that for some type of analyses, it may not be necessary to transcribe the entire source; 'selected sentences, passages, paragraphs, or stories relevant to the research question or theory may be all that are needed' (Emerson, Fretz and Shaw, 1995, cited in McLellan, MacQueen and Neidig, 2003). Indeed, the method adopted in this study included the use of both keywords and paraphrasing as well as literal transcription of sections of discussion. More specifically, when a concept, which was fundamental for the analysis (e.g., because linked to one of the *a priori* codes), was not expressed in a particularly sharp or impactful way, and/or when it was expressed through quite wordy and lengthy sentences, keywords and paraphrasing was used to save both time and the essence of

the concept. When, instead, a concept or idea was expressed in a meaningful and “emotive” way, quotes were annotated (with juror code and time of the discussion) to be then reported verbatim in the thesis. The approach described resulted, in line with previous literature, in suggestions to use the tapes as a supplement when the analysis is not aimed at providing in-depth accounts of a given phenomenon on individual and/or collective level, but at the exploration and identification of general themes and patterns, whereby the investigation can be carried out with less detailed text (McLellan, MacQueen and Neidig, 2003).

To further clarify and to conclude on this point, the choice made in this context and the reasons provided for it are not intended to disregard the utility of full transcripts, and were put forward notwithstanding an appreciation of the fact that full transcripts were successfully used in studies similar to the one presented here. However, as McLellan, MacQueen and Neidig (2003) pointed out, when taking this sort of decision, researchers’ evaluations have to be made through a consideration of, among other things, cost/time required to accomplish the task. Past research projects, carried out by research teams, used various methods of transcription, including transcription of the deliberations on-the-fly (e.g. Dutwin’s, 2003) or, when such financial costs could not be afforded, trained transcribers to embark on a several-months-long “transcription journey”. Neither of these solutions were plausible for the present study, given the individual nature of PhD research as well as consequent limited time and financial resources.

The second view of the video deliberations and the activity of partial transcription provided an opportunity to begin identifying new emerging, *a posteriori* themes as well as to amend, specify, and/or re-categorise the *a priori* ones. Therefore, at the end of the second view, frequencies were counted and number of occurrences of phenomena annotated. This is not a “strictly” qualitative approach, yet it is often used in content analyses in that they are conceived as always virtually quantitative as well as qualitative (Leedy and Ormrod, 2013). While this aspect will be more specifically addressed later, it is worth specifying here that numbers were annotated with the intention of having a grasp on how frequently an identified phenomenon occurred,

which also enabled comparisons and highlighted differences between conditions; in so doing, there was no intention, however, to use this specific dataset for quantitative statistical analysis.

4.3.5 Stage 5: Video deliberations – third view: double-check with questionnaires

As suggested by research methods literature, especially in cases when full transcripts are not used, the audio-/video-tapes should be replayed as needed (Stewart and Shamdasani, 2014). This is equivalent to the usual rereading of written data, which is an essential element of improving the rigour of qualitative analysis and the coding process (Renner and Taylor-Powell, 2003). Therefore, in order to explore that source of data once again, video deliberations were watched for a third time, in an effort to further validate the relevance and systematic organisation of both *a priori* and *a posteriori* codes and themes as well as checking the accuracy of recorded frequencies. The further goal of this third view was to use the videos and questionnaires concurrently, checking the latter against the former. This extra step, first and foremost, ensured higher rigour of the data analysis by ascertaining the accuracy of the data collected, and this allowed consideration of whether data reported from the questionnaires matched data deduced from the videos (e.g. jurors' individual verdict preferences pre-/post-deliberation), and, where there was a mismatch, to double-check whether that was due to an error in the transcription or the participant's claim not corresponding with their written response. Moreover, for questions that concerned jurors' perceptions of the situation (e.g. the presence of leaders in the group), re-watching the videos and consulting the questionnaires led to further reflections on discrepancies between jurors' perceptions and the actual unfolding of deliberation dynamics.

4.3.6 Stage 6: Video deliberations – final view and collection of data from questionnaires

Video deliberations were watched one last time with no specific purpose other than further increasing the validity and accuracy of the material collected for the analysis.

To this end, an all-comprehensive final check was conducted while watching the videos. Questionnaires were used again – this time separately from the videos – to collect information needed for the statistical analysis. Indeed, the questionnaires, given the nature of some of the questions posed (a mix of Yes/No, open and scaled responses), produced quantitative data, which enabled answers to some of the research questions as well as to benefit the whole study with the advantages of triangulation. Therefore, starting from the related research questions, relevant variables were identified to then be tested through quantitative analyses.

4.3.7 Qualitative Analysis

Consistent with the highly structured nature of the employed template analysis technique (section 4.3.1), the following analysis was oriented towards a systematic and organised approach. Therefore, a qualitative content analysis of video deliberations and text of the deliberations' transcript was undertaken. 'A content analysis is a detailed and systematic examination of the contents of a particular body of material for the purpose of identifying patterns, themes, or biases' (Leedy and Ormrod, 2013, p.148). According to Leedy and Ormrod (2013), this type of analysis is usually performed on 'forms of human communication, including [...] videotapes of human interactions, transcript of conversations...' (p.148). As Forman and Damschroder (2007) pointed out, content analyses can be different - quantitative or qualitative – while sharing the central feature of the systematic categorization of data. For this reason, since tabulating frequencies of recurring elements is a fundamental step in any content analyses, some would argue that such analyses are always quantitative as well as qualitative (Leedy and Ormrod, 2013). This somewhat "hybrid" connotation can be found also in the present study: the step of counting frequencies was taken at the outset, in order to identify recurring themes, patterns and biases and, most of all to compare the number of occurrences between the two different experimental conditions. The analysis conducted can be, therefore, defined as a 'qualitative summative content analysis', which, following Hsieh and Shannon's (2005) classification, 'involves counting and comparisons usually of keywords or content, followed by the interpretation of the underlying context' (p.1277). Subsequently, on

the basis of what resulted from the counting, an analysis of the manifest and latent meanings has been conducted and explanations for the identified phenomena have been proposed.

4.3.8 Quantitative Analysis

Along with a qualitative examination of the data collected, descriptive and inferential statistics were carried out on the data that was quantitative in nature. Indeed, for reasons that were already elsewhere specified (Chapter 2), a mixed method approach was employed in this thesis, in order to appropriately answer the various research questions and sub-questions posed by the research problem under analysis. “IBM SPSS Statistics 24” software was used to run relevant statistical analyses, both descriptive and inferential. Quantitative data for statistical analyses were gathered from the pre-/post-deliberation questionnaires, which generated a mix of categorical, open and scaled responses, hence numerical data suitable for statistical analyses (Punch, 1998). As anticipated, the questionnaires were specifically aimed to acquire information about mock jurors’ individual views on the deliberations and their outcomes, so that their otherwise undetectable personal preferences would still emerge, even if the group discussion tended, voluntarily or involuntarily, to bury them. Details about findings and potential explanations for them will be provided towards the end of this chapter (section 4.4.2), where an account of the specific statistical analyses conducted will be provided along with related comments on the findings and their triangulation with qualitative results.

4.4 Findings and Discussion

Findings from this study will be reported and discussed simultaneously in this section. Results obtained through the previously described analysis of qualitative and quantitative data are interpreted, with the aim of finding explanations for the recurrence of certain phenomena in (mock) jury deliberations. Particular attention will be given to the two different conditions in order to reflect on the impact of the variables under analysis: effects of presence/absence of a (mock) judge and of presence/absence of the requirement for motivated verdicts are investigated to

answer the posed research questions (Chapter 2, section 2.3). The section will address separately the qualitative and quantitative findings.

4.4.1 Mock Juries' Deliberations – Qualitative Content Analysis

Through the employment of the described template analysis technique, relevant *a priori* and *a posteriori* themes were identified, throughout the different stages of the investigation, across the data qualitatively analysed. Table 4.1 shows the coding template as resulted after the final stage of analysis.

Table 4.1: Qualitative Content Analysis – Coding Template

THEMES	CODES	FURTHER CODES	DESCRIPTION
Jurors' Behaviour	<u>Approach to deliberation</u>	N/A	Juries decisions on how to start (vote first, discuss first, etc.)
	<u>Narrative Construction</u>	N/A	Jurors' tendency to make sense of evidence creating stories
	<u>Confirmation Bias</u>	N/A	Jurors' tendency to start from their own beliefs and find explanations in the evidence, rather than the opposite
	<u>Leadership</u>	N/A	Emergence, within juries, of leading personalities
Errors	<u>Memory Errors</u>	N/A	Errors jurors make when failing to remember facts of the case
	<u>Evidence-related Errors</u>	<i>Misunderstanding/ Misconception</i> <i>Overestimation</i>	Errors jurors make when failing to understand evidence and/or its probative value
	<u>Law-related Errors</u>	<i>Misunderstanding of Reasonable Doubt</i>	Errors jurors make when failing to understand legal matters
Differences (between conditions)	<u>Presence of the mock Judge</u>	<i>Advantages:</i> <ul style="list-style-type: none"> – <i>Summary Report</i> – <i>Questions</i> – <i>Clarifications</i> – <i>Accuracy</i> 	Advantages of having a professional on jury panels
		<i>Disadvantages:</i> <ul style="list-style-type: none"> – <i>Frequent interruptions</i> – <i>Influencing remarks</i> – <i>(Too) Early opinion disclosure</i> – <i>Control over topics addressed</i> 	Disadvantages of having a professional on jury panels

4.4.1.1 Jurors' Behaviour

The first broad theme identified was labelled “Jurors’ behaviour” and encompasses aspects that describe how jurors generally behave, reason and act throughout deliberation. As indicated in Table 4.1, the theme included four specific codes, which are individually addressed in the following sub-sections.

4.4.1.1.1 Approach to deliberation

The first aspect that was observed were the preferences that jurors exhibited towards voting at the outset of the deliberation, that is once the hypothetical crime case had been described, the pre-deliberation questionnaire completed, and the mock juries left by the researcher inside their fictional “deliberation room”. All the mock juries were left with instructions on deliberation and verdict: they were all specifically asked to reach a unanimous verdict and, in addition, the five British mock juries were expressly told they could freely choose how to conduct the deliberation. This last instruction was not given to the five Italian mock juries, where the mock judge had, amongst other duties, that of directing the discussion (according to the information acquired through the judges’ interviews – see Chapter 3). The five Italian mock juries’ deliberations, thus, consistently started with a “go-round” procedure, which was requested by the mock judge, who before expressing her own opinion, asked each lay member to declare what their vote was and to explain why, according to the Italian jury system’s rules (CPP, 2010). This was intended to get the discussion started, prompting the jurors to reflect on the motivations underlying their initial opinion of guilt or innocence, more than asking them to simply vote.

While the presence of a mock judge granted consistency in the choice of deliberation approach across all the Italian groups, the five British mock juries proceeded in different ways. Details about each group’s decision were recorded and are briefly summarised below:

B1: J3 suggested a roundtable vote and all the other members agreed; despite the agreement, the group did not complete the first poll, as the discussion naturally started;

B2: the group did not take a specific decision on how to start, they began discussing the issues with jurors randomly intervening and expressing their opinions/votes;

B3: the group started with a first poll; J3 asked the other members who thought the defendant was guilty and who thought he was not guilty, and jurors voted by a show of hands;

B4: J2 asked the group whether they wanted to start with a vote; J3 replied that she thought they should instead start with a discussion of the evidence. However, the group ended up voting beforehand.

B5: J6 explained – as if there were a set rule – that they should go through and each of them should briefly state their opinions; following, when other members tried to motivate their first vote, he stopped them, claiming they should only vote at that stage. The group, therefore, ended up voting beforehand.

Previous empirical studies have found a fairly even split between two different approaches to deliberation: taking a vote at the outset or discussing beforehand and voting at the end (Ellsworth, 1989). More recent studies have highlighted the inappropriateness of such a clear-cut distinction between the so-called “verdict-driven” and “evidence-driven” approaches, showing that in most instances the approach actually chosen resulted in a mix between the two (Ellison and Munro, 2010a). This indeed happened in the present study, where, as described above, even if a specific approach was suggested, it was not always followed.

Due to the needs of experimental conditions, only five groups were left free to choose their approach to deliberation, and the small number of groups and variety of choices do not allow highly reliable empirical conclusions on whether or not the choice of approach has had an impact on deliberations. However, the variety and inconsistency of approaches chosen, along with the intentions shown, represent interesting points of consideration. It may be worth noticing that, regardless of what actually happened, the vast majority of groups (all except B2) tended to suggest voting before discussing. Interestingly, it has been argued that rational decision-making, in groups’ face-to-face

discussions, should be supported by a process that moves in the opposite direction. It should, in fact, start with information exchange, move on to a presentation of underlying arguments (and expected resolution of conflicts), and close with an actual decision-making stage wherein decisions are taken and votes are expressed. 'Anything else is considered a deviation from the "best" procedure, resulting in poorer quality decisions' (Whitworth and McQueen, 2003, p.2).

Hastie, Penrod and Pennington (1983) explained this in the specific context of jury decision-making, pointing out that, when starting with expressing one's vote, jurors feel bound to that vote and focus their deliberative efforts on defending that position rather than keeping an open-mind towards facts of the case, evidence and issues that may come up during the discussion. This may result in attention shifting away from the real problem and being focussed instead on endless, heated discussions driven by the individual's need to prove their point. Previous studies have observed the effects of various approaches on deliberations and seem to confirm Hastie, Penrod and Pennington's (1983) claims. Ellsworth (1989), for instance, found that juries that voted later deliberated for longer, and spent more time discussing issues actually relevant to the case. In addition, Ellison and Munro (2010a) observed that juries that tended to embark on heated and competitive discussions were those that opted for early verdict polling.

If what has been observed so far is true, jurors who find themselves in deliberative contexts wherein votes are taken early may tend to focus on defending their positions in heated debates, rather than prioritising the interests of the defendant and the fairness of the trial. In such instances, decision-making will still be centred on reasons behind choices, as this thesis fundamentally assumes; yet, those reasons will not necessarily be the most (legally) relevant ones. By contrast, when initial preferences are not expressed and a formal vote is postponed, jurors will be free to pay attention to the relevant issues, silently reconsider their position and, in some cases, privately change their mind. In light of the presented arguments, there is reason to believe that, even if not to be conceived as dichotomous, the choice of one of the various approaches to begin deliberating is not without consequences. This should perhaps

lead to reflection on whether juries need to operate under set voting procedure instructions, in order to grant greater consistency to this stage of the decisional process.

4.4.1.1.2 Narrative Construction

After observing the groups' initial choice of how to approach the deliberation, the analysis focussed on the individual deliberative dimension as a fundamental component of the collective decision-making task juries must undertake. Indeed, jury deliberations offer an interesting opportunity to reflect on collective group decision-making dynamics and the effects of the interaction among a number of decision-makers having to operate towards a common goal. However, the understanding of any complex phenomenon as a whole has to be grounded in a preliminary grasp of the functioning of its components. Likewise, to fully understand jury decision-making, it is crucial to understand juror decision-making first. To achieve this, jurors' deliberating behaviour was scrutinised in the context of the present study, in an effort to identify cognitive processes individual jurors undertake and recurring patterns among jurors' behaviours in the mock jury deliberations conducted.

Research has long been proposing explanations of the cognitive processes that lead jurors throughout their decision-making task. Well-established theoretical models have mostly resolved the matter through two types of approaches: mathematical-based and explanation-based (Hastie, 1994). A more comprehensive discussion on this multifaceted theoretical framework is provided in the literature review chapter of this thesis (Chapter 1) and therefore a detailed analysis of the various positions will not be reiterated here. Suffice it to say, for the purpose of this part of the discussion, that mathematical-based approaches posit that jurors conduct an assessment of the defendant's culpability based on calculations of value attributed to each piece of evidence, while explanation-based approaches suggest that jurors elaborate a cognitive organisation of the evidence in order to make sense of it and, in so doing, create a story that reconstructs what they think has happened (Winter and Greene, 2007).

The idea of jurors as passive recipients of pieces of information, which they would then weigh and evaluate to reach a conclusion on the basis of a probabilistic calculations, despite potentially being desirable, does not appear to reflect the reality reported by jurors (Devine et al., 2001; Ellison and Munro, 2010a). Therefore, explanation-based models – most of all, the story model (Pennington and Hastie, 1988; 1993) – are preferred as theoretical explanations that better reflect the reality of what happens in the minds of the jurors. Jurors, as active thinkers, with their own past knowledge, preconceptions and expectations, try to make sense of the facts of the case, assembling the evidence into a story that provides a coherent causal explanation which fits their past experience, knowledge, etc. (Hastie, Penrod, and Pennington, 1983; Hastie, 1994).

One of the most evident problems with the fact that jurors do not enter the jury box with their mind as a *tabula rasa* (Kirgis, 2002), but rather with their own background knowledge which plays a role in how they will make sense of the evidence presented, is that jurors will all hear the same evidence, and yet they may construct different stories. The way narratives are constructed will determine how the evidence is weighted and interpreted and that, in turn, will determine the juror's final decision (Hastie, 1993). The dangerousness of this process is also intuitively understandable; in constructing a narrative, jurors will prioritise the evidence that best fits their own story and/or will make the interpretation of evidence serve the story, so that the evidence that does not fit may be disregarded or interpreted in a different way (Pennington and Hastie, 1988; Lempert, 1991). However, this process clearly does not lead to fairly evaluated evidence, when considering that evidence presented at trial should be assessed on the basis of its actual probative value. Whether or not lay jurors are capable of such an evaluation is a different issue, which will be addressed in detail later. The problem here is that narrative construction, as a decision-making process, does not lead to an assessment of evidence probative value, even if that value is correctly understood.

Consistent with the theoretical propositions of the story model, the findings of the present study demonstrated that most mock jurors indeed tended to create narratives

to make sense of the evidence in the context of what they perceived to be a plausible reconstruction of the events. Through the conducted content analysis, frequencies of the occurrence of this phenomenon were recorded in order to obtain a general idea of the recurrence of narrative construction and draw conclusions on the incidence of this jurors tendency. Moreover, the phenomenon was measured in the two observed conditions (British and Italian) separately, in order to see whether it occurred more frequently in one condition than in the other and, if so, try to understand why. It was hypothesised that, of the two variables under analysis, the presence of a (mock) judge in the Italian condition may well have an impact on this phenomenon. Indeed, given the risks that a narrative-based evidence interpretation may bring to the process, Italian judges who sit on juries claim to take the responsibility of inviting jurors to focus on the evidence's objective probative value, whenever jurors attempt to explain their position by creating a scenario of how they think things went. This piece of information, also acquired at an earlier stage of the research project through judges' interviews, was implemented in the experiments, so that the mock judge acted like the real judges in similar occurrences.

Analysis identified that the creation of narratives across all ten groups occurred forty-four times; thirty times in the British condition and fourteen in the Italian condition. The typology of stories was most disparate, varying from very imaginative to more plausible reconstructions.

J3 B1: *'My theory behind this whole thing was: the neighbour wanted to go around, he saw as a good opportunity that she's broken up with her boyfriend - I know I'm just assuming but this is the story - he's got a good opportunity as she's broken up with her boyfriend, he's gone around with a bottle of wine, thinking something could happen maybe after... She said "No, I don't wanna have sex with you", he maybe, potentially raped her and then he's gotten aggressive and killed her.'*

This quote clearly exemplifies how narratives are constructed and brought to the discussion. It is particularly interesting to notice how evidence is barely mentioned here and the scenario is reconstructed entirely on the basis of speculation (as the juror seems to realise). The only piece of evidence that is actually mentioned (and not even considered with reference to its probative value) is a bottle of wine, which, according

to the description of the case had the defendant's and victim's DNA on its rim (Appendix G). However, that is not to say that narratives which included multiple bits of evidence constituted a guarantee of greater accuracy or more sensible, legally-oriented reasoning. For example:

J3 B3: *'I think he (the defendant) is not guilty because she was raped, and they found a condom with the guy's semen and she was fully clothed and there was a wine bottle with both DNA on it, so they could just been having a nice drink, did what they had to do, she got fully clothed and there was a gun [...]'*

Moreover, in sixteen cases, the stories included another individual – the victim's ex-boyfriend, who was mentioned in the description of the case (Appendices E, F, G) showing that jurors, in their need to assign blame, were willing to fall into the trap set by the Defence who tried, in the closing arguments, to shift the attention onto the alleged (and quite weakly proven) presence of the ex-boyfriend at the crime scene. Therefore, even when the mock jurors did not have enough evidence to convict the defendant, they tended to explain that in light of the culpability of the ex-boyfriend.

J3 I3: *'The two guys (defendant and ex-boyfriend) surely had a fight because one took the gun off the hand of the other.'*

J2 I5: *'The intercourse was consensual, the ex gets there and they fight because the defendant gets jealous.'*

J2 B2: *'[...] That makes me wonder if it started off consensual, she decided not to and he carried on... I think that's what's happened. On top of that, I mean, the only thing I can really think of, logically, why this has happened is that in fact someone else, maybe Frank (the ex-boyfriend) is coming, seeing them having sex, he's trying to stop him... which explains the fingerprints on the barrel [...]'*

J3 B3: *'I think maybe the neighbour had a gun trying to scare off the boyfriend because the boyfriend had the other gun and he was like: "Shoot her!" and then she shot him and he shot him... it makes sense!'*

The reported quotes, besides highlighting quite strongly how imaginative and bizarre some jurors' reconstructions may be, also show that the tendency to create narratives

was present in both the British and Italian conditions. The above-reported difference in number of occurrences, however, cannot be disregarded and may not be coincidental. Jurors in both conditions were all jury eligible laypeople and there is no reason to believe that they would go through different cognitive decision-making processes. In fact, embracing the story model of jury decision-making, it has to be concluded that all mock jurors in this experiment underwent similar cognitive processes and, accordingly, should have shown roughly the same tendency to the creation of narratives. In the face of these observations, the most reasonable explanation for this difference has to be attributed to the different experimental conditions, in particular to the presence of a judge.

In light of the findings, it appears that the presence of a judge had an impact on the phenomenon of story creation in that the mock judge, following the instructions received, stopped jurors from carrying on with representations of scenarios whenever they attempted to use this “method” in order to prove their point. For instance:

Mock Judge (to J3 I1): *‘Rather than hypothesise what he (the defendant) may have done, I would urge you to consider what has been found at the crime scene and how that has to be interpreted for the purpose of a conclusive judgement of guilt or innocence, beyond any reasonable doubt... Making assumptions and speculating about events we didn’t witness is not the correct logical/legal path to take in order to make this sort of decision.’*

The very fact that the mock judge had to intervene with such clarifications offers further evidence of the reflection presented earlier about the plausibility of the story model approach working in both conditions: Italian mock jurors tended to make sense of the case through narratives as well as British mock jurors. It is sensible to observe though that the judge’s repeated warnings, even if directed from time to time to individual jurors, were listened to by the entire jury, which led other jurors to avoid repeating the error. This, as observed in the data, did not avoid the reiteration of these types of errors altogether; Italian mock jurors still tended occasionally to create narratives, even when the judges’ warning had already been communicated. The presence and contribution of the mock judge, thus, can only account for a reduction in the occurrence of this behaviour amongst jurors.

The presented findings and observations are not intended to imply that the presence of professionals on a jury panel would eliminate the root of the issues. There will remain problems with the creation of stories on the basis of different, potentially arbitrary, interpretation and weighting of evidence in light of individual juror's background knowledge, experience and beliefs. In fact, the present analysis could only take into account those "stories" that mock jurors decided to tell aloud. It might be the case, thus, that some other procedure should be in place to grant a more "legally-oriented" deliberation. One of these could be, as will be seen later, the requirement for motivated verdicts. If jurors tend to bring into the deliberative processes elements – such as all the "ingredients" of the stories – that are not legally appropriate, asking them to focus on legal reasons for their choices could be a further aid to use when the presence of a judge is not enough.

4.4.1.1.3 Confirmation Bias

As a corollary of the jurors' tendency to make sense of the evidence through the construction of narratives on the basis of previous knowledge, experience and beliefs (Pennington and Hastie 1986, 1988, 1994), a further related tendency was identified: mock jurors' frequently conducted reasoning that started from their own conclusions and worked backwards to explanations of the evidence. These instances, *a posteriori* coded under the label "Confirmation Bias", were easily recognisable and distinguishable from the opposite cases, because it was clear from the words of the mock jurors that the "direction" of the reasoning was from a given conclusion backwards. For example:

J6 B1: *'I think he (the defendant) raped her.'*

J5 B1: *'She was redressed...'*

J6 B1: *'Sometimes people redress their victims as a sign of remorse, so if he genuinely cares about her and has got this kind of infatuation with her, he probably would feel quite guilty...'*

Regardless of whether the specific observation made is reasonable or not, what is relevant here is how the explanation is not offered as a reason underlying the conclusion drawn (the defendant raped her); it is, instead, found only afterwards when it becomes necessary to justify the conclusion against the objection made (the victim

was fully clothed, hence the hypothesis of rape is unlikely). This is perhaps even more evident in the following example:

J5 B1: *'If you do use a condom, then you take care of getting rid of it, I guess.'*

J6 B1 (starting from the same belief as above): *'Well, but if you just shot someone maybe you don't have time to get rid of the condom, because you think someone could have heard.'*

This claim, which follows the same schema as above, appears even more meaningful in light of the fact that it was the same juror (J6 B1) that a few minutes earlier declared it strange that: *'no one has heard the gunshot'*. Thus, the juror knew that, according to the evidence presented, there was no witness who admitted to have heard the gunshot, yet the same juror was willing to disregard this element in order to justify the conclusion from which they started.

In other cases, the occurrence of this particular type of cognitive process was made even more evident by the fact that the explanation of evidence that followed the need to justify a previously rooted conclusion was quite objectively unrealistic.

J1 B2: *'I believe he (the defendant) is guilty mostly because his testimony contradicts pretty much all the evidence at the crime scene [...] Regarding the fingerprints, I understand it's weird that they are on the barrel and not on the trigger, but it's possible that one cleans the gun and doesn't think to clean the barrel too.'*

Finally, another demonstration of the inverse reasoning process was found in the fact that, at times, the same element was explained by two different mock jurors in opposite ways, in accordance to their own previous ideas and conclusions. Interesting in this sense are the two following claims made by mock jurors with regard to the defendant's bruising:

J2 B6 (convinced that the defendant is not guilty): *'Given his previous convictions, I think it's possible that the bruising has nothing to do with the crimes and depend on the brawl.'*

J5 I1 (convinced that the defendant is guilty): *'You see it from the previous convictions, he is an aggressive violent individual, that's typical of sexual violence.'*

These findings confirm the claim that people reason bi-directionally (Glöckner and Engel, 2013). Specifically, it has been argued that when people are faced with the task of making decisions based on multiple judgements and inferences, this process is not unidirectional – flowing from the evidence to the conclusion – but it is in fact bidirectional, that is ‘the evidence influences the conclusions and, at the same time, the emerging conclusion affects the evaluation of the evidence’ (Simon, Snow and Read, 2004, p.814). According to this view, algebraic models for integration of evidence have to be rejected as plausible theoretical explanation of the functioning of jury decision-making. Indeed their assumptions that pieces of evidence are evaluated and specifically weighed independently in the process of reaching a final rational and informed decision are incompatible with the bidirectional nature of an actual juror’s reasoning, as shown by previous research and by the current findings. On the contrary, this phenomenon could be theoretically explained by so-called parallel constraint satisfaction models, which posit that when decision-makers attempt to produce coherent mental reconstructions of a situation, they automatically and involuntarily change their perception of the evidence (Holyoak and Simon 1999; Thagard and Millgram 1995; Glöckner and Betsch, 2008). Consequently, as was observed in the experiments, the information that supports the decision-maker’s preferred interpretation of the facts is emphasised and given more importance, whilst the information that would be against such an interpretation is given less relevance or neglected altogether (Nickerson, 1998; Glöckner and Engel, 2013).

In terms of implications, it has to be noted that this way of reasoning is the opposite of what should happen inside the deliberation room, because, through this cognitive process, the original information is systematically distorted, which may produce further consequences (Holyoak and Simon 1999; Simon et al., 2001; Simon, Snow and Read, 2004). If a juror has come to believe that the defendant is guilty/innocent on the basis of inherently ambiguous evidence (as that presented in the hypothetical crime case used in the experiments), these opinions may well bias the juror’s interpretation of the evidence (Hendry, Shaffer and Peacock, 1989). Moreover, Simon, Snow and Read (2004) argue that this way of information distortion is liable to increase people’s confidence in their (perhaps inaccurate) ideas and decisions, which would most likely

happen because the very information that could destabilise confidence (i.e. the conflicting evidence) is detracted in value and discarded. It could be argued that the deliberation process may counteract this phenomenon, in that the discussion of different opinions may lead group reasoning to outperform individual reasoning. However, previous research has demonstrated that an individual's confirmation bias, prompts people, when they reason alone or with peers, to potentially strengthen their original beliefs (Mercier and Landemore, 2012). Therefore, as well as narrative construction, confirmation bias and any cognitive process that hinders an objective and legally-oriented evaluation of the facts acquires worrying connotations and should be controlled.

4.4.1.1.4 Leadership

The last code identified under the broad theme "Behaviour" is leadership. Since dynamics of leadership are inherent to all groups, these emerged from the video deliberations and, as such, were analysed. While it would fall beyond the scope of this thesis to extensively address forms, causes and consequences of leadership, focus will be on understanding whether such group dynamics have developed in different ways in the two different experimental conditions created for this study. It was expected that these dynamics would develop differently depending on the lay or mixed composition of the jury panel, in that the different *status* of the professional judge could not be disregarded as a factor that would have a significant, conscious or unconscious, influence on leadership processes during deliberations (Zambuto, 2016).

When considering the figures of judges and jurors alone, several distinctive characteristics emerge: judges, in any jurisdiction, belong to a different, narrower population (legal professionals) than jurors (members of the general public); judges have access to case information and possess the knowledge and expertise that is necessary to fully understand it, while jurors do not. Judges have a *forma mentis* that leads them to bring legal motives and values to their decisions, as opposed to jurors who, because of naivety and/or inexperience, may deliberate on less legally-oriented motives and values (Champagne and Nagel, 1982). All these (and other) characteristics

are brought into deliberations and certainly play a role in determining how leadership works in lay jury panels and mixed jury panels.

In fact, in all panel judicial procedures, according to the Italian Criminal Justice System, there is a predetermined leader, who is officially recognised in the person of the “president”, a judge who is nominated to direct deliberations with other judges. In the Italian trial by jury, this role is granted to two judges (“*giudici togati*”), who operate with lay members on jury panels. The leadership of these individuals is recognised in both types of procedure, but obviously even more strongly in jury trials, where judges are seen as members provided with decisive authority, especially by jurors who, lacking the same degree of experience, cannot disengage themselves from their dependence on a judge’s aid (Gulotta, 1987). The fact that, by contrast, the British trial by jury does not impose a predetermined leader leaves space for leadership/influence dynamics to operate naturally within jury panels and has generated differences that mock jury experiments were able to observe.

Leadership has been observed from two different perspectives for the purpose of this thesis. The video deliberations provided qualitative insights whilst post-deliberation questionnaires offered data for quantitative analysis and for further triangulation. To first address the qualitative analysis, it is worth specifying that a detailed description of observed within-group leadership dynamics will not be provided for the Italian mock juries. Across these groups, affected by the presence of the mock judge as a predetermined leader, leadership dynamics were mostly consistent, with the mock judge playing a distinctive leading role. However, differences were found across the British mock juries, wherein the lack of a predetermined leader left room for interesting leadership dynamics.

The five British mock juries all behaved in different ways. Interestingly, in all groups neither a foreperson was elected nor did anyone volunteer to lead the discussion. This contributed to the natural emergence of different personalities and demeanours during the unfolding of the discussions. Therefore, in B1 and B3 there were no explicit leading tendencies during the discussion of contrasting arguments, conducted among

members who did not tend to prevail or overpower the others. However, heated debates arose in B2, where two mock jurors (J1 and J3) quite evidently showed strong leading tendencies. Their fervent defence of conflicting verdict preferences resulted in the two jurors monopolising a great deal of the overall deliberation, which at times turned into a two-sided “fight” that the other members watched, occasionally showing consent or dissent by nodding or shaking heads.

J3: *‘The fingerprints don’t say anything.’*

J1 (sarcastically): *‘Don’t say anything. DNA evidence at the crime scene doesn’t say ANYTHING...’* (continuing looking at J3 with eyes wide open and appearing stumped): *‘Do you wanna literally have watched him shoot her in order to... no? I think I don’t understand what more you could want...’*

And again, later on in the discussion:

J3: *‘For example, she wants to kill her ex-boyfriend because they have financial problems, so she said to the neighbour “Oh, do you have a gun?” – “Yes” – “Can I see the gun? Can you show me the gun?” Ok, we can discuss about that.’*

J1: *‘Right... I feel you are just, like, making irrational leaps to, like, filling in gaps to create a scenario that is not the most likely. In this case you go with Occam’s Razor: the most likely scenario is almost always what happened. It’s probably NOT her trying to buy a gun off her creepy neighbour.’*

And, lastly, considering that the rest of the group seemed to support J3’s verdict preference (not guilty):

J1: *‘I really don’t understand why all the evidence is to the contrary of the victim wanting anything to do with her neighbour and now people are willing to throw it all away for...’*

J1 eventually accepted the “not guilty” verdict for the purpose of reaching the required unanimity, as indicated by rating the verdict “extremely unfair” in her post-deliberation questionnaire. She also added, in a personal conversation with the researcher immediately after the experiment:

J1: *‘If this was real, it’d have been a hung jury, because there is no way I’d have let them acquit him!’*

With the remaining groups, it was interesting to notice how the presence of two leading, contrasting personalities can be managed differently. Two leading personalities (J2 and J3) emerged in B4 as well, although this did not result in particularly heated debates. Although it was clear that they had contrasting opinions, this was always conveyed in a reasonable and peaceful way:

J1: *'At this point, given the inconclusive evidence, we cannot convict him for rape either.'*

J3: *'I don't think I can agree he is not guilty of anything because I'm not even sure he didn't kill her.'*

J1: *'Well, I would still rather let a guilty man free than an innocent man in jail.'*

Quite different from all the others was the situation in B5, wherein one leading personality immediately emerged. As reported earlier, J6 began explaining confidently (albeit inaccurately) how the group should proceed with the first ballot. He used the same confident and quite confrontational demeanour throughout the entire deliberation, even when putting forward unfounded arguments (as it will be seen in the following section). In reaction to this behaviour, there were some initial attempts to constrain him and find space to voice a perspective:

J2: *'So, for me the relevant evidence is, 1) bruising on the eye and cheek...'*

J6: *'There is no evidence he (the defendant) did that...'*

J2: *'Can I just talk and then...?'*

In response to this, J6 left some space for the other members to express their opinions, however he still spoke much more than the others and did not show openness towards the opinions of others. It was quite clear that he had his mind made up from the outset and was not willing to change it. This behaviour seemed to progressively exert an effect on the other participants, who initially contradicted him, but were later convinced by his arguments and accepted his verdict preference even when it did not correspond to theirs.

4.4.1.2 Errors

Further focus of attention, in the context of the present qualitative content analysis, was the commission of errors on the part of jurors. Given the widespread criticism of the functioning of the jury system (Devlin, 1965; Kassin and Wrightsman, 1988; Brooks, 2004), the occurrence of errors throughout these mock deliberations came as no surprise and was, in fact, pre-emptively identified as one of the broad *a priori* themes that were expected to emerge. The codes used at the outset of the analysis of this theme were, as the template shows (see Table 4.1, “Memory Errors”, “Evidence-related Errors”, “Law-related Errors”). These three widely defined categories of errors were expected to occur, given that previous literature has extensively demonstrated the fallibility of human memory and its repercussions on the criminal justice system (Loftus, 1975) as well as the existence of serious difficulties and failures that laypeople face in attempting to understand complex evidence and legal principles (Lanza, 1994).

4.4.1.2.1 Memory Errors

To begin with memory errors and their effects on the criminal justice system, one of the most investigated areas in this context is the unreliability of eyewitness testimony due to its dependence on human memory (Loftus, 1975). Although eyewitness testimony itself was not a particularly relevant issue in this study, the conclusions drawn in this area of research in terms of fallibility of human memory may be adapted to the jury situation as well. This is because witnessing a crime and being involved in a criminal trial (as a juror) are experiences which share some characteristics: they are both stressful, they both require attention, they both expose individuals to quite lengthy memory retention intervals, and so forth. Indeed, metaphorically, jurors themselves can be considered witnesses of an event: the trial. Hence, their memory is very likely to lead to similar inaccuracies (Salerno and Diamond, 2010).

Extensive research has shown that human memory is fallible and that errors can occur at any time during the acquisition, retention and retrieval of information regarding a witnessed event (Loftus, 1996a). These errors obviously affect jurors’ cognitive processes as well as any other individual, and even more so considering the situation they are experiencing. First of all, broadly speaking, information may be lost at the

acquisition stage due to factors that might be connected to the event itself (in this case, the trial) or to the witness of the event (in this case, the juror). Event factors can be, for example, time-related factors (e.g. the length of the testimony, which can span over several days, and/or the delay between hearing the testimony/evidence and partaking in the deliberation); but also location condition; opportunities to focus on details; salience of details. Witness (juror) factors regard mainly the perceptual activity that individuals engage in, e.g. to what details they paid attention; previous expectations; the level of stress that the individual experiences. Both event and witness factors are very likely to affect jurors' memories during the encoding of information released at a trial, so that a specific piece of information might be missed and not encoded at all.

The following examples show cases in which mock juror memory issues occurred at either the acquisition or retention stage.

J1 I5: *'The semen in the vagina was determined to be his (the defendant)? Didn't remember that, to be honest.'*

J3 I3: *'I'm not sure if there were two guns or one.'*

J2 I3: *'I didn't manage to catch all the elements, the details, the evidence... I'm confused.'*

While there were, among the examples above, cases in which the individuals were aware that their memory of the event was flawed because they missed something while the event was happening, that is not always the case. When there is no realisation the error occurred at this first stage, that particular piece of information is lost and will not be included in the subsequent reasoning, that is, for jurors, in the decision-making process.

However, even when the information is correctly perceived, errors might occur as a result of transformations in the retention phase, during which information may be forgotten or subject to interference, and therefore, eventually misremembered (Ainsworth, 1998). In a jury context, this may be caused, for example, by inaccurate

information supplied by group members, which can influence other group members' memories (Salerno and Diamond, 2010). The original information can be contaminated in various ways by post-event information (Loftus, 1996a). For instance, it is possible that post-event information affects original memories through the insertion of new details that contradict the original ones (Loftus, 1979; Loftus and Loftus, 1980; Loftus, 1996a; 1996b; Christiaansen, Sweeney and Ochalek, 1983 ; Zaragoza, Belli and Payment, 2007). Examples of these were found in the current study:

J2 B4: *'On one gun there were fingerprints of Piggott (the defendant) and nobody told us where the fingerprints were, and on the second gun there were fingerprints on the barrel [...]*

J3 B4: *'I think "gun 1" was proved to be the murder weapon.'*

(Original information: there were two guns; on the barrel of one gun there were the defendant's fingerprints; nothing was said about evidence on the other gun; neither of the two guns were conclusively proved to be the murder weapon).

J3 I3: *'DNA on the bottle is the ex-boyfriend's, not his (the defendant's).'*

(Original information: DNA from both the victim and the defendant was found on the rim of a wine bottle at the crime scene).

Post-event information can also alter original memories by causing a rearrangement of existing elements, or even introducing non-existent elements, creating in both cases false memories (Loftus, 2000):

J3 B3: *'They found semen around her vagina.'*

(Original information: semen was found inside the victim's vagina)

J6 B5: *'She's got bruising on her neck.'*

(Original information: there was bruising in the victim's genital area)

Lastly, when errors did not occur during the first two phases, they can still affect information that is correctly perceived and retained, but compromised by internal or external factors that make it inaccessible for final recall. Internal factors are

represented by stereotypes or scripts, that is background knowledge or prefigured conceptions of typical events, which may create distorted memories (García-Bajos and Migueles, 2003). External factors are usually found to be the result of misleading questions, which may induce a certain response and, in so doing, significantly affect the memory retrieval (Hall, Loftus and Tousignant, 1984). Although internal factors, as a result of internal processes, are difficult to detect, it seems likely that there is room for those factors to intervene, and for related errors to occur in jury settings, as they are based on the same assumptions on which the story model of jury decision-making is grounded. On the other hand, examples of external factors were found in the data:

J3 B3: *'The (ex)boyfriend had scratches on his back, didn't he?'*

J3 B4: *'Didn't the prosecution say that the first gun was proved to be the murder weapon?'*

So far errors that occur at the various stages of an individual's memory functioning have been considered and how they can generally occur in a jury context has been exemplified. However, another important category of memory errors, which are likely to occur in jury contexts because of the peculiarity of the situation, are the so-called inferential errors. Any criminal trial is permeated by a degree of uncertainty, which leads the decision-makers to base at least part of their decision on inferences and deductions. Accordingly, it is reasonable to believe that jury deliberations constitute a breeding ground for the proliferation of this specific type of error; the inferential errors are memory errors that result from inferences that people make and that can be mistaken for actual memories (Hannigan and Tippens Reinitz, 2001). For example, during the description of the fictional criminal case a piece of information given was that fingerprints of thumb and forefinger of a left hand were found on the barrel of the gun, and that those fingerprints were the defendant's. Starting from that information, mock jurors said the following:

J3 B3: *'I remember about the (ex)boyfriend how he is left-handed and the fingerprints on the gun was something about left-handed...'*

J4 I4: *'He (the defendant) isn't left-handed, so it's difficult to believe that he shot the gun... The fingerprints on the gun were from a left-hand.'*

As these examples clearly show, starting from pre-existent beliefs, these mock jurors inferred elements (i.e. the handedness of defendant and ex-boyfriend), thinking they were remembering them. Interestingly, J3 B3 had declared that she thought that the defendant was not guilty and that the ex-boyfriend was involved in the crimes; thus, the mock juror's memory was 'adapted' accordingly: the ex-boyfriend did it, fingerprints were from a left hand, the boyfriend was left-handed. Moreover, the second example shows how the belief of the defendant's innocence can be supported by fabricated memories as well (even without blaming the ex-boyfriend): the defendant is not guilty, fingerprints were from a left hand, the defendant is not left-handed; this was the inferential process of J4 I4. As previous research already demonstrates, inferential processing is able to influence jurors' memories (Holyoak and Simon, 1999). In reality, none of the above-reported inferences were supported by the evidence actually presented, and in fact this inferential processing carried out by jurors affected their memories to the extent of compromising their reconstruction of the facts (e.g. it was never mentioned whether the defendant, and even more so the – not formally accused of any crimes – victim's ex-boyfriend, were right or left-handed).

The issues considered so far predominantly relate to the individual dimension of memory functioning and errors. Of relevance to this discussion is also its collective dimension, that is how these cognitive processes function within groups. Some would argue that groups' performances are generally better in this respect, giving rise to fewer memory errors than individuals (Clark, Stephenson and Rutter, 1986; Vollrath et al., 1989; Hinz, 1990; Maki, Weigold and Arellano, 2008). An explanation for this allegedly better performance would lie in the fact that groups should be able to catch memory errors of which individuals might not acknowledge the occurrence. In fact, this is one of the reasons why criminal justice systems rely on jury deliberations to determine someone's guilt or innocence (Ruva, McEvoy and Bryant, 2007). However, this view is challenged by contrasting empirical research, which has found that groups might be even more susceptible to memory errors. Individual jurors are not sensitive to the accuracy of their memories, hence their incorrect recalls can influence other

group members' memories (Salerno and Diamond, 2010), which in turn creates limitations in the aptitude of deliberations to generate memory improvements (Pritchard and Keenan, 2002).

This study, albeit not specifically comparing individuals and groups performances in terms of memory errors, offered an opportunity to observe how groups reacted to individuals' memory errors. This, in turn, might leave room for consideration of the beneficial or detrimental effects of deliberations on jurors' memory and memory errors. Most importantly, the present study produced a further element of analysis, in that groups' reactions to memory errors were different also depending on the experimental condition (British or Italian), and more precisely, on the presence/absence of a (mock) judge on the panel.

The mock judge, present in the Italian condition was given a written summary report, which she read aloud at the outset of each deliberation, and which constituted a source of information available for her consultation at any time, to dispel potential doubts on the facts of the case. This experimental/procedural difference impacted on the occurrence of and reaction to memory errors within groups, which altogether resulted in the following findings: occurrences of memory errors were recorded 18 times across groups in the British condition, and only 6 times across groups in the Italian condition. This difference in the findings is unlikely to be coincidental, and can instead reasonably be seen as a result of the effects of the mock judge's presence/absence. Effects which, through observation and analysis of the video-deliberations, can be pinpointed as impacting deliberations in the following ways.

First of all, the presence of a mock judge prevented some memory errors from occurring, given that mock jurors in the Italian condition heard the crucial elements of the case one extra time through the reading of the summary report, as opposed to mock jurors in the British condition, who only heard the description of the case during the trial stimulus. Moreover, there was the possibility for Italian mock jurors, when they did not remember something, to ask the judge rather than guessing. Both these aspects, by aiding mock jurors' memories, influenced the decision-making task, and

generated observable effects which highlighted differences between the two experimental conditions. These aspects will in fact be further addressed later in the discussion, as the judge's use of a summary report (and its effects on deliberation) constituted one of the main differences between the two experimental conditions and, as such, fell within a different category in the thematic hierarchy (see section 4.4.1.3). However, it is important to mention it here, as this certainly had an effect on preventing memory errors from occurring in the Italian condition deliberations.

While the effects of the mock judge's presence described so far intervened pre-emptively avoiding the occurrence of certain memory errors, this did not eliminate memory errors entirely. Indeed, and perhaps more alarmingly, not even the presence of the mock judge, with the reading of the summary report and the possibility to ask questions, were able to completely avoid memory errors on the part of the mock jurors. Italian mock jurors, going through similar cognitive processes as British jurors ended up making similar mistakes. Nonetheless, effects of the different procedural conditions could still be observed in the way memory errors were dealt with in the two different experimental situations. British mock jurors' memory errors had to be dealt with in the context of a panel of peers, where everyone was at the same level and everyone's opinion had the same weight. The absence of a directive figure (e.g. a judge) and/or of a source of information to consult (e.g. a summary report), which could provide *super partes* answers, in order to easily resolve memory gaps or inaccuracies, resulted in different outcomes. On the one hand, there was the risk of "transferring" errors:

J2 B4: *'There were fingerprints on the gun...'*

J3 B4: *'But whose fingerprints? That hasn't been said.'*

J2 B4: *'Yeah, I don't think so.'*

(Original information: the defendant's fingerprints were found on one gun).

On the other hand, and by contrast, there could be disagreement about elements that were remembered differently by different mock jurors and this led to (sometimes long and heated) debates, which often ended with an inaccurate conclusion that was, nonetheless, accepted by the group. Going back to the example of the mistaken

memory of the defendant being left-handed, which was arbitrarily fabricated only because the fingerprints found on the gun were from a left hand, it is interesting to notice how the same error was dealt with very differently in the two conditions. Starting with the British condition:

J3 B3: *'... Because apparently he was left-handed... Was he left-handed?'*

J4 B3: *'I'm not sure we were told that.'*

J3 B3: *'Yeah, we were.'*

J1 B3: *'The second gun had a thumb from the left hand...'*

J3 B3: *'Yeah, from the left hand - there we go - and the guy was left-handed.'*

By contrast, in the Italian condition:

J4 I4: *'He (the defendant) isn't left-handed [... as reported above]'*

Mock Judge: *'Hold on, sorry, we need to be careful here: we only know that the fingerprints found on a gun at the crime scene are from the defendant's thumb and forefinger of his left-hand... But we don't know if he is left- or right-handed. We weren't told that.'*

In this second example, the mock judge, armed with authority and accurate information, has rapidly resolved an issue that, in the British mock jury, was not resolved, and led to an inaccurate conclusion. That inaccurate conclusion was carried on throughout deliberation and affected consequent thoughts as well. It has to be taken into account that such situations certainly occur in real-life jury trials and with even greater incidence, when considering that actual deliberations last much longer than experimental ones and that 'the likelihood of inferential errors increases substantially as the retention interval increases' (Hannigan and Tippens Reinitz, 2001, p.939). In light of these observations and bearing in mind the seriousness of the consequences of mistakes in real life scenarios, rethinking the management of jurors' memory errors appears necessary.

4.4.1.2.2 Evidence-related Errors

Another broad category of errors that tend to be made during deliberations are evidence-related errors. For the purpose of this thesis, evidence-related errors are defined as all errors that jurors tend to make when (mis)interpreting, (mis)conceiving,

(over)estimating evidence. To use the eyewitness testimony example, research has extensively shown that jurors tend to misunderstand the way memory works and, therefore, overestimate the reliability of eyewitness testimony to an extent that has frequently led to wrongful convictions (Schmechel et al., 2006). Yet, this is only one area where such errors usually occur. Broadly speaking, the deep-seated scepticism towards jurors as decision-makers finds its foundations in the belief that these laypeople generally do not possess the necessary “equipment” to properly understand complicated issues and evidence, because of the amount, complexity and specificity of the information on which they should base their decisions (Shuman and Champagne, 1997). Accordingly, one of the strongest criticisms against juries’ capabilities to adequately understand and assess evidence regards scientific evidence, of which jurors often appear to incorrectly estimate the probative value (Imwinkelried, 1983). This has led to arguments around juries being ‘incapable of adequately understanding evidence or determining issues of fact, [...] they are unpredictable, quixotic, and little better than a roll of a dice’ (Shuman and Champagne, 1997, p.255).

The present study, consistent with previous literature, provides examples of the occurrence of a wide range of evidence-related juror errors. This phenomenon, which was reasonably foreseen (and therefore identified as one of the *a priori* themes), was subject to further scrutiny, from which, additional sub-categories of evidence-related errors emerged. As Table 4.1 shows, errors related to evidence were grouped into two subcategories: misunderstanding/misconception and overestimation. For the purpose of the present discussion, ‘misunderstanding/misconception’ of evidence refers to incidents when jurors fail to accurately understand a piece of evidence – in these cases, jurors tend to attribute improper meanings to evidence presented and/or to develop incorrect, arbitrary ideas about what the evidence is actually revealing about the facts of the case (Schklar and Diamond, 1999). ‘Overestimation’ of evidence refers to those cases in which jurors, even if properly understanding the evidence, tend to attribute to it too high a probative value, by perceiving weak evidence as strong and decisive elements of their decision.

A relatively high occurrence of these sorts of errors occurred in the mock jury experiments – a total of 42 occurrences in the British condition and 14 in the Italian condition. The total, for both conditions, includes errors of both types and reveals the existence of some commonalities amongst jurors' evidence-related errors, both within and across conditions. This suggests that, regardless of the specific group's influence or any procedural difference, there are some common grounds to the commission of certain errors in interpreting and estimating evidence and its probative value. Overall, data showed that the most common mistakes occurred regarding the weakest (i.e. the most ambiguous) pieces of evidence in the hypothetical trial scenario.

Previous empirical research, specifically focusing on jurors' perceptions, understandings and assessments of evidence (especially forensic evidence, mainly DNA), has found it difficult to qualify evidence as "strong" or "weak" for experimental purposes. This methodological pitfall derived from the lack of objective measures to judge evidence strength (Anderson, Schum and Twining, 2005). This issue, albeit acknowledged, was not specifically addressed for the purpose of the present study, which was not designed to capture such a complexity in evidence strength or weakness. Therefore, in the context of this study, the evidence that is defined as "weak" is quite obviously weak, in that it is undoubtedly (at least, in the eye of a well-informed observer) overly ambiguous and inadequate to support any conclusion in favour of the defendant's culpability.

It is even more meaningful that, despite the obviously weak nature of most evidence used in the experimental stimuli, participants in the present study still committed errors in the interpretation, understanding and evaluation of the weakest evidence. Indeed, the weaker the evidence is, the more likely it is, given the lack of clear and objective situational factors, that jurors will misinterpret it (Kaplan and Miller, 1978; Kassin and Wrightsman, 1988; Kovera, McAuliff and Hebert, 1999). By contrast, according to Ruva, McEvoy and Bryant (2007), strong evidence makes situational factors clearer and less subject to interpretation, leading in turn to individual juror differences becoming almost irrelevant to the decision. This phenomenon is described as the Liberation Hypothesis (Kalven and Zeisel, 1966; Devine, et al., 2009; Bjerregaard

et al., 2017), and postulates that the strength of the facts become a constraint for jurors, who cannot interpret them “freely”; jurors are, therefore, “liberated” and free to employ the whole set of their background knowledge, bias and beliefs when the evidence is weak or ambiguous.

I. Misunderstanding/Misconception

In the mock jury deliberations conducted for this study, one of the most common evidence-related errors was the misunderstanding/misconception of the presence of a condom at the crime scene containing the defendant’s seminal fluid (Appendices E, F, G). The condom evidence could be used and interpreted in several contrasting ways by someone who did not immediately realise that its own ambiguity should have deprived it of any strength or, at least, reduced its probative value. Amongst the multiple explanations that were given for the presence of a condom as an element against or in favour of the defendant’s culpability, the most frequent and inaccurate inference referred to the widespread belief that a rapist would not use a condom. Precisely, this inaccurate and unfounded remark was made 20 times by mock jurors across all the groups. For example:

J4 B1: *‘Would a rapist wear a condom?’*

J2 B1: *‘That’s what I was wondering...’*

J5 B2: *‘If you wanna rape someone, really... I don’t think you use a condom.’*

J2 B3: *‘There was a condom used, which... I’ve never heard that in a rape case.’*

J1 B4: *‘The fact that he used the condom is inconsistent with rape.’*

J5 B6: *‘He (the defendant) used a condom...’*

J6 B6: *‘Yeah, do rapists do that?’*

J4 I3: *‘They don’t use condoms in such crime, these are impulsive crimes.’*

J2 I4: *‘The use of a condom is impossible during a rape.’*

These few examples taken from the video deliberations provide a clear picture of how misconceptions and inaccurate previous beliefs are brought into deliberation by jurors. In fact, the use of a condom on the part of sexual offenders has been long acknowledged by the academic literature and, given its newsworthy nature, more widely diffused through newspaper articles (which are presumably a more accessible source for laypeople). The New York Times reported that condom rape cases were recorded, by the San Francisco Rape Treatment Center, to be between 15% and 20%, in 1994 (Wolff, 1994). Likewise, early academic literature recognised the occurrence of such a phenomenon and found various reasons for perpetrators' choices: rapists might use condoms to avoid leaving evidence of semen in the victim's body, but also as a way to protect themselves from diseases, if not as a ritual in serial crimes (da Luz and Weckerly, 1993; Blackledge, 1996). This precautionary behaviour appeared to become more common as awareness of DNA identification and sexually transmitted diseases increased (Blackledge and Vincenti, 1994; Maynard et al., 2001; Grubin, Kelly and Brunsdon, 2001). More recent research has also confirmed that, while sex offenders' condom use might have been rare in the past (Woodhams and Labuschagne, 2012), the phenomenon has increasingly been reported and the reasons are still mainly connected to heightened awareness of the use of DNA evidence in criminal investigations (Spencer et al., 2011; Bradshaw et al., 2013).

When comparing these data with the ostentatious confidence of some of the mock jurors' remarks reported above, it is easy to understand the situation and its worrisome consequences. These types of errors, seen through the lens of theoretical narrative models of jury decision making, result in misconceptions that are connected to jurors' overreliance on their misinformed background beliefs (Pennington and Hastie, 1986; Wagenaar, van Koppen and Crombag, 1993), which may be even more pronounced in this case due to the highly resistant (and pervasive) nature of rape myths (Ellison and Munro, 2010b). Alarming, jurors who are wrongly convinced of a belief that does not find any confirmation in reality (but perhaps only in their schemas and bias about how sexual offences are perpetrated) may base their verdict decision on these beliefs.

Moreover, similar errors and misunderstandings occurred in relation to the gun/s. The presence of a second gun at the crime scene (Appendices E, F, G) was not given the emphasis that it should have had in weakening the Prosecution's contention that the defendant had used the murder weapon to kill the victim. One of the most problematic misunderstandings in this respect was that the gun on which the defendant's fingerprints were found was very often uncritically and implicitly considered the murder weapon, although mock jurors were told that no conclusive ballistic results had confirmed that.

J5 B2: *'It was said that that (the first gun) is the murder weapon.'*

J2 B4: *'There were fingerprints on the gun...'*

J1 B2: *'... But his fingerprints on the murder weapon at all should be indicative of him being in contact at some point with the murder weapon at the scene of the crime, near the time of death, and there is no other explanation for that contact.'*

J3 B2: *'I'm surprised the Prosecution didn't say anything about the other gun...'*

J2 B2: *'So it's assumed that there is nothing relevant on it for them not to bring it up.'*

J4 B3: *'I personally think the presence of the two guns, one with his (the defendant's) fingerprints on it, is the most important bit of evidence in favour of conviction.'*

II. Overestimation

The other sub-category of evidence-related errors, which emerged from the data gathered included those cases in which the probative value of evidence was overestimated. This generally happened when jurors attributed to one or more piece/s of evidence greater strength than it actually deserved in proving what in fact happened. Throughout the mock juries' deliberations this happened several times and related to various matters. One of the most meaningful examples regarded the involvement of the victim's ex-boyfriend (who, despite not being accused, was very often mentioned as an alternative culprit) and, in particular, the high probative value

that mock jurors frequently attributed to the evidence, that placed the ex-boyfriend at the crime scene (Appendices E, F, G).

J2 B2: *'I think it was the ex-boyfriend...'*

J1 B2: *'I agree... They had financial problems and his bank statement was found at the crime scene.'*

J3 B2 to J4 B2 (who was the only one at that point that considered the defendant guilty): *'Are you swayed? Because that is hard evidence!'*

Obviously, the presence at the crime scene of a bank statement of a person who until a week before the crime/s had been in a long-term relationship with the victim, and therefore quite likely visited the apartment on a regular basis, is not 'hard evidence' at all. Neither can one consider strong evidence (at least not on its own) a testimony, or lack thereof, that placed the ex-boyfriend at the apartment:

J4 B2: *'If there was an eyewitness who had placed the ex-boyfriend there, I would have changed my mind, but because there is none, but the neighbour (the defendant) was definitely there, for me that's enough, combined with the gun.'*

A final, quite meaningful example, concerns the accusations of rape. Different to the murder – which clearly happened as the victim was dead – regarding the rape accusation, it was for the mock jurors to determine whether the victim had been raped or not, before determining whether the defendant committed that crime (or not). This raised even more discussions and doubts, and at times errors of the overestimation of the evidence's probative value. For example:

J2 B4: *'How do you explain the signs of rape?'*

The description of the fictional case specified that evidence found during an examination of the victim was consistent with potential sexual violence or violent sexual intercourse, therefore evidence of rape was inconclusive and the sexual intercourse may have been consensual. Regarding consent, those mock jurors who reasonably considered the matter still made errors.

J5 B4: *'I don't believe there was rape... There was a positive relationship: she let him in.'*

In this case, for example, the weight attributed to the existence of a positive relationship between the defendant and the victim appears to be inappropriate in light of the inferences drawn. The implication of this quote is that the victim consented to sex just because she let the defendant into the flat, which may be attributed to rape myths/misconceptions. Indeed, the above-mentioned reasoning ignores the possibility that consent initially given may have later been withdrawn, in which case rape must be considered to have occurred (Lyon, 2004). Yet, to continue on the matter of consent, it is interesting to note how errors of this sort could also be found in those who correctly considered the hypothesis of consent being given and later withdrawn. An example can be taken from group I2, where J2 insisted that she found the defendant guilty of rape in that she believed there was proof that consent was given and then retracted. The mock judge asked J2 I2 to explain the reasons why, in light of the evidence, she thought there was proof of consent being withdrawn.

J2 I2: *'... Even the fact that there was semen both in the condom and in the vagina...'*

Mock Judge: *'This gives you proof of what exactly?'*

J2 I2: *'Of the fact that it was initially consensual intercourse and then it became sexual violence. They started with the condom and carried on without.'*

Mock Judge: *'Actually, no. This gives you proof that there was sexual intercourse. Period. Also, we only know whose semen is in the condom (the defendant's), but not whose semen was in the vagina. Why do you take it for granted that that's necessarily the defendant's?'*

J2 I1: *'I don't know. That's the idea I had.'*

Judge: *'Ok, but eventually ideas and culpability judgements have to be anchored to the evidence.'*

When considering that this mock juror, notwithstanding the reasonable observations put forward from the mock judge, maintained her views until the end of the deliberation and then with difficulty agreed to accept the not guilty verdict proposed by the group, it appears clear how powerful these sorts of errors can be in real-life criminal cases. Mock jurors' overestimations of evidence as well as other evidence-

related errors create conditions whereby jurors carry their own – inaccurate – evaluation of the evidence’s probative value and base on that their final decision.

In addition, another aspect that this last example highlights is the fact that the judge’s clarification offers an opportunity to remind jurors (at least the ones who are willing to be reasonable decision-makers), firstly, of aspects they are disregarding (in this case, unknown identity of semen in vagina) and secondly, of the fact that they must remain focussed on the evidence, for the purpose of a fair decision. The mock judge exerted this role throughout all deliberations and therefore clarified points that were sources of errors as exemplified above: specifically, with reference to the above-mentioned examples, the mock judge clarified that the gun was never proven to be the murder weapon and, likewise, specified that the ex-boyfriend was not the defendant and that the bank statement could have been there legitimately, given the existence of a romantic relationship between the victim and him. These, along with other clarifications provided by the mock judge, explain the difference in numbers of errors observed in the Italian and British conditions. For example:

J4 I3: ‘... *But were there fingerprints on the other gun?*’

Mock Judge: ‘*No. Not that we know of.*’

This clarification dispelled any potential doubts regarding the presence of fingerprints on the other gun and, therefore, errors on this aspect did not occur. Clarifications of this sort demonstrate how the occurrence of some errors was in fact promptly prevented by the intervention of the mock judge in the Italian condition, whilst it could not be avoided in the British condition.

4.4.1.2.3 Law-related Errors

Another category of errors that jurors tend to make are law-related errors, which are those errors that occur due to a lack of knowledge of legal matters. Because jurors’ lack of legal expertise is well-known (Frank, 1973; Elwork, Sales and Alfini, 1977; Vidmar, 1994; Fisher, 2000) and often used as an argument against their ability to fairly deliberate, these errors are usually expected. They may concern any legal principle and the knowledge and/or understanding of it. Throughout the analysis

conducted for this study, the more serious and diffuse legal misconception regarded the reasonable doubt standard of proof, which was widely mentioned and poorly understood.

I. Misunderstanding of Reasonable Doubt

Among the variety of law-related errors that lay jurors make, given their lack of expertise and knowledge of the legal system (Lanza, 1994), particular attention should be given to the widespread misunderstanding of the legal principle: “reasonable doubt”. Incorrect conclusions were often drawn by mock juries with regard to the required standard of proof whereby an individual can be convicted only if considered “guilty beyond a reasonable doubt” (Dhami, Lundrigan and Mueller-Johnson, 2015). The observance of the reasonable doubt standard is one of the core principles that lies at the foundations of the jury system. Pennington and Hastie (1981) included it in the main components of the overall jury task, showing that complying with this standard is a step of the jury decision-making task that cannot be overlooked. However, despite being a fundamental legal principle, the reasonable doubt standard of proof has always encountered serious issues in its interpretation and application: ‘the only accepted, explicit yardstick for reaching a just verdict in a criminal trial—is obscure, incoherent, and muddled’ (Laudan, 2003, p.295). There is great conceptual confusion, which leads to philosophical debates of difficult resolution even for those (often not laypeople) who ask themselves, first and foremost, ‘what is “reasonable”?’.

Thinking critically about reasonable doubt does not appear to be an activity to which jurors are generally prone. As it has been observed in previous studies and confirmed by the findings of the present, ‘not one person on any jury raised the question of the definition of reasonable doubt’ (Ellsworth, 1989, p.221). This issue becomes even more alarming when considering that the standard of reasonable doubt is frequently brought into the discussion by jurors themselves. According to Ellsworth (1989), for example, mention of this standard of proof is often made when more lenient factions cannot convince harsher members to change verdict preference through arguments regarding facts and evidence, and therefore they decide to obtain the result using reasonable doubt as a persuasive instrument. These findings are confirmed by the

results of the present study. In four out of five British mock juries, reasonable doubt was not mentioned at the outset of the discussion as a standard to take into account, but it was only mentioned a few minutes before the end of deliberations, in an attempt to convince those who still considered the defendant guilty. Only in B5, reasonable doubt was (incorrectly) introduced at the outset by J6.

J6 B5: *'To find someone guilty, you have to be 100% certain.'*

These observations, consistent with findings of previous research (Ellison and Munro, 2010a), lead to the conclusion that the reasonable doubt standard is left to be invoked by jurors who should use it as a threshold, and in fact do not even question the intrinsic meaning of the concept. As a result, jurors who believe they know what reasonable doubt is (merely because it sounds familiar), trust their idea that in order to find someone guilty beyond a reasonable doubt one must be 100% certain of that person's guilt (Laudan, 2003). Consistently with this, the standard "100%" was used to define "reasonable doubt" in all British groups except for B2.

J6 B1: *'There is not enough evidence to say it was 100% him.'*

J2 B3: *'I'm not going to change my mind because there is no way I can say 100% he did it.'*

J2 B4: *'We can't say 100%...'*

J6 B5: *'Even if you are 99% sure that someone did it, you cannot convict him, you got to be 100% certain.'*

This belief in a "100% certainty" does not leave room for any "reasonability" of the doubt in question, since it would rather equate to "no doubts whatsoever". Such a standard would lead to the paradoxical conclusion that almost no one could be considered guilty of a crime. Within the criminal justice system, the "legal truth" very rarely has the opportunity to be checked for consistency with the "actual truth" (Summers, 1999), because – with exception of very few cases – almost no doubt can be 100% dispelled. As a consequence, it is not true that jurors must be 100% certain of

guilt in order to convict, because 'it is not true that any doubt is a reasonable doubt' (Laudan, 2003, p.316).

The resolution of such problems is not easy. A number of solutions have been proposed, but do not satisfactorily resolve the question. For instance, leaving judges the task of instructing jurors in this respect (as already happens in some legal jurisdictions) does not solve the problem completely; judges as well as jurors have different and discordant interpretations of the notion (Laudan, 2003). Some would argue that, given the philosophical intricacies embedded in the concept itself, attempting to subjectively explain it would make it even more ambiguous: 'I find it rather unsettling that we are using a formulation that we believe will become less clear the more we explain it' (Newman, 1993 - Chief Judge of the U.S. Court of Appeals). By contrast, others, convinced that the problem lies in the subjective nature of the interpretation of a vague concept, have suggested that quantification of the notion is the only effective and legally feasible solution to adopt (Saunders, 2005).

Despite the benefit that adding some degree of objective clarity to the interpretation of the concept undoubtedly presents, quantifying it leaves room for another set of problems that, ultimately, make it unadvisable to follow that route. One of the most compelling critiques of this approach came directly from judges, some of whom sensibly observed that 'percentages or probabilities simply cannot encompass all the factors, tangible and intangible, in determining guilt – evidence cannot be evaluated in such terms' (Simon and Mahan, 1971, p.329). A similar argument was also raised more recently, when it was argued that it is not appropriate to apply a "grid system" approach to cases (Bree, 2007 cited in Gunby, Carline and Beynon, 2013).

In other words, it appears impossible to quantify the concept of reasonable doubt, especially using percentages or probabilities. There are indeed significant doubts about the capability of jurors to understand probabilities when they are used in trials with reference, for example, to DNA evidence (Kaye and Koehler, 1991; Nance and Morris, 2005). Consequently, it is difficult to see how this could be different regarding standards of proof. However, even if that limitation was overcome and jurors were

able to understand and use percentages and probability, the problem would remain in depriving the concept of a necessary degree of subjectivity, which would be against the very nature of the jury system.

Taking into account all considered thus far, it has to be concluded that the notion of reasonable doubt, as it stands now, is so vague to become rather meaningless (Saunders, 2005). As Lauden (2003) compellingly argues, justice cannot be granted in systems where it is officially accepted that juries apply discrepant, subjective standards to determine guilt or innocence; there is no way to exclude, in such systems, that two juries faced with the same case will decide differently because they have differently interpreted the standard by which they must abide. Such systems lack uniformity and predictability and, for this reason, are intrinsically unjust. To attempt to find a solution and to turn the reasonable doubt standard into a meaningful construct, the concept has to be provided with 'a tangible meaning that is capable of being understood by those who are required to apply it' (Blackmun, 1994, p.29). In this respect, what emerged from this study reinforces such views.

To compare procedural differences between the two experimental conditions and to observe potential consequences, mock jurors in the British condition were left free to deliberate as they wished, whilst Italian mock jurors were required to follow some rules. Most of those rules were connected to the presence of the judge who, amongst other things, at the beginning of each deliberation set out the basic principles of the Italian criminal procedure. The mock judge warned the mock jurors that in making their decision they had to take into account three legal principles: the presumption of innocence, the burden of proof on the prosecution, and the reasonable doubt standard. The principles were explained and mock jurors were asked whether they understood. The fact that all Italian mock jurors claimed they understood was not always confirmed in the discussions, since some of their remarks were contradictory. The issue was, nevertheless, promptly overcome each time through the mock judge's intervention, which further clarified or simply reminded the mock jurors of the principle and of its meaning. With specific regard to the reasonable doubt principle, the following examples show the effectiveness of the mock judge's intervention:

'You don't have to necessarily be convinced that the defendant is innocent, but you might not be sufficiently convinced that he is guilty.'

'There will always be a remote possibility that things didn't go that way and therefore we'll never be 100% certain, we'll always have a minimum doubt, right? Now, having a doubt is NOT quite the same as having a reasonable doubt.'

'Our "certainty" about guilt has to be anchored to the evidentiary framework. On the basis of the evidence, of all the evidence together considered, I must have a firm belief that the defendant is guilty. If all the evidence, conjunctly and coherently evaluated, do not convince me that the defendant committed the crime, I must acquit.'

In these cases there were errors and claims from jurors that evidenced misunderstanding of the concept, which were clarified and corrected every time by the judge. Accordingly, what happened in the Italian mock juries seems to be in line with the only sensible conclusion that can be drawn, that is, that a certain degree of objectiveness in the interpretation of the meaning of reasonable doubt should be introduced. As others have already observed, and as the findings of this study confirmed, in order to encourage jurors to more thoroughly scrutinise the evidence, there is probably a need for 'a fairly straightforward instruction aimed at neutralizing the bias' (Simon, 2004, p.519). In order to find a balance between this need for objectivity and the contrasting need to leave the jurors some discretionary power, so as to avoid depriving the jury trial of its own inherent democratic powers, the emphasis has to be placed on the evidence. It is not really a matter of the juror's state of mind, it is rather a question about the concrete, observable relationship between known events (i.e. evidence) and strength of the support that it can provide to a theory of guilt (Lauden, 2003).

4.4.1.3 Differences between the two conditions

A further phenomenon that was foreseen and identified as an *a priori* theme was: "Differences between the two conditions". As shown in Table 4.1, this included the code "Presence of the mock judge". The occurrence of differences between the British and the Italian conditions was hypothesised at the earliest stages of the present research project, which was designed to measure the effects of these differences.

Consequently, the overarching theme was expected to include also aspects of difference regarding the other variable under analysis in this thesis, namely the presence/absence of a requirement for motivated verdicts. However, while differences ascribable to the presence of a (mock) judge were mostly notable through behaviours and dynamics in the video deliberations, differences resulting from the requirement for motivation were mainly found among the motivation-related responses gathered from the post-deliberation questionnaires. Accordingly, the former generated qualitative data for the present content analysis, whereas the latter were subjected to quantitative analysis and will be reported later in the related section (4.4.2.9).

4.4.1.3.1 Presence of the mock judge

Within the code “Presence of mock judge”, two further codes were identified throughout the analysis: “Advantages”, which encompasses all the positive aspects of the presence of a mock judge, as emerged at various stages of the deliberations and through comparison between conditions, and “Disadvantages” which includes all the problematic aspects that the presence of a mock judge generated, and that emerged through the observation of the mock judge behaviour and the mock jurors reactions throughout deliberations.

Before addressing the differences in more detail, it is worth clarifying that, since the diversity of the two systems is at the heart of this research, some of the following differences will have been mentioned elsewhere in the thesis, specifically where each of them appears relevant to the phenomenon in analysis. Notwithstanding, for the sake of clarity, it seems useful to propose here a comprehensive view of these differences as a further, more structured source of information.

I. Advantages

– Summary report

As previously noted, all the Italian mock juries’ deliberations began with the judge reading aloud a summary report, which provided a recap of the most relevant information in the case. This activity, and other aspects of the mock judge’s behaviour,

were planned and designed in accordance with the real judges' accounts regarding their roles (Chapter 3). As expected and as emphasised in Chapter 3, the summary report helped jurors better understand and remember fundamental elements of relevance to their decision and had an impact on the entire decision-making task. Moreover, the introduction of a summary report in the experiments enabled observation, also through comparison with the British condition which lacked this report, of the effects of the employment of this tool.

The reading of the report constituted a fundamental aid for Italian mock jurors' memories. As already observed, these mock jurors had greater access to the relevant information, which was repeated to them one extra time before deliberation. Extensive research has demonstrated that repetition improves memory (Popov and Reder, 2017), and well encoded elements are less confused (Kılıç et al., 2017), hence, the additional exposure of those mock jurors to the core information allowed them to better encode and thus better remember it. This was further increased by the observed tendency of mock jurors in the Italian condition to continue, during the reading of the summary report, to take notes and double-check those that they had already taken (an opportunity that mock jurors in the British condition did not have).

– Questions

In some cases, this double-checking activity resulted in actual questions asked to the judge during the reading and/or in the request to reread a particular section of the report which was not clear:

J3 I1: *'... So, were fingerprints and semen both the defendant's?'*

J1 I4: *'Excuse me, could you please read again this section? I think I missed something...'*

The opportunity to ask questions to the judge was not only in response to memory flaws and to the summary report, as questions were asked, throughout, also when some points were not clear at later stages of deliberation. Given the presence of a

judge on Italian jury panels, questions could easily be asked at any point. As already specified, to grant experimental consistency across groups, all mock jurors (British and Italian) were warned that they could not ask the researcher any questions on the case and they had to deliberate only on the basis of the information provided during the description of the case. Nevertheless, mirroring what happens in reality, Italian mock jurors could ask the mock judge any questions at the outset and throughout deliberation. British mock jurors, who could not promptly refer to a professional to dispel their doubts, ended up either asking the other jurors (who might be equally uncertain of something) or silently resolving their doubts, introducing uncertainty and inaccurate ideas throughout deliberation, and letting them affect their final decision.

– Clarifications

Further help for Italian jurors might come from clarifications the judge felt necessary to make. As identified in Chapter 3, judges are aware of the fact that lay jurors are not used to the legal system and might find it difficult to fully understand its functioning. Taking this into account, they tended to shed light on matters that they know from experience could be a source of misunderstandings.

A first example of this was found from the outset of deliberation, when the mock judge, even if it was not specifically indicated by the script of the summary report, took care in specifying who the various individuals in the case were. For example, when the written report indicated the defendant's full name, the mock judge read it and additionally specified his role in the trial:

'Mr Piggott... who is the defendant in this trial, [...].'

Along the same lines, the mock judge also always made it clear that:

'The defendant is only one, the counts of indictment are two.'

Such clarifications were frequently necessary. Indeed, confusion often arose regarding correct identification of the defendant and of the counts of indictment. Specifically, very often the victim's ex-boyfriend, despite not being charged with any crime, was

treated by the mock juries as if he were the defendant. Furthermore, the two counts of indictment were not always given the same relevance and attention, when lay jury members, left on their own, embarked on disorganised discussions wherein reference to one or the other crime was randomly made.

Whether this sort of confusion is based on the fact that jurors actually forget, or do not understand, or simply momentarily disregard the relevant elements of their decision, the problem remains. While the mock judge's precautions show the existence of a worry towards likely confusion, it is certainly beneficial here to remember that this issue was also raised by the Italian judges interviewed for the first study (e.g., Judge 1 and Judge 3, Chapter 3, pp.93-94). Judge 1's claims constitute further confirmation that this issue deserves attention, as it may well be one further element in the development of inappropriate decision-making processes that may lead to unfair decisions.

– Accuracy

Since a degree of uncertainty is inherent to all criminal trials and to criminal justice proceedings in general, there is often no known objective truth against which the "fairness" of deliberation outcomes (i.e. verdicts) can be checked. Nonetheless, regardless of the final decision and its perceived fairness, any jury deliberation could be conducted in a more or less accurate manner. This depends on a series of elements: maintaining focus on legally relevant factors, anchoring opinions to the evidence presented at trial and avoiding misunderstandings which permeate the entire discussion. Many of the arguments presented earlier in this discussion demonstrate that the intervention of the mock judge ensured greater accuracy to the deliberations (e.g. greater compliance with the procedures/processes/law). The mock judge ensured that the discussion was equally focused on both counts of indictment, that it took into account legally relevant factors, that evidence was correctly remembered, and that misunderstandings would be quickly resolved. Some examples of the mock judge's intervention in this respect included:

'Okay, we discussed the murder already. Now we are focusing on the rape...'

'You shouldn't use the evidence to create an imaginary dynamic of the events.'

'Why do you think so? ... But, careful, I mean on the basis of what has been found, not of what you think might have happened.'

'We don't know which of the two guns was the murder weapon.'

II. Disadvantages

Some of the difficulties that purely lay juries encounter are probably reduced in jury decision-making when there is a judge present, who is familiar with legal principles and all that is needed to make a well-informed and legally fair decision. That is not to say, however, that those difficulties are completely warded off if there is a judge (judges are, after all, also fallible human beings), but it is assumed that risks of unfair, arbitrary decisions are reduced. A further problem remains, since other elements of arbitrariness can occur throughout the process that leads to the final decision. Some of these factors regard the very presence of judges on jury panels. Considerations in this respect have already been made in the context of Chapter 3, and findings from this study further support those arguments.

Judges who operate within the Italian criminal justice system, even when required to decide collegially, are used to working with other professionals, so that in such instances a group of experts decide delicate and often quite intricate criminal matters. In cases when a mixed composition of the judicial panel (e.g. Italian juries) is required, judges are in a peculiar position for which they do not receive instructions or training. They, therefore, decide to moderate their behaviour on the basis of a few, quite broad indications that the Code of Criminal Procedure gives them. As a result, although the Code seems to propose a mechanism oriented to avoid judges influencing lay members, that appeared to be almost the opposite of what happens (Chapter 3).

Some of these influential tendencies emerged throughout the Italian mock juries' deliberations and, accordingly, lower order codes were developed to account for the recurring problematic behaviours that were evident across the various Italian groups.

– Frequent interruptions

The mock judge frequently interrupted lay members during the discussions and, in so doing, she prevented them from freely express their thoughts, which generated misunderstandings. For example:

J5 I2: *'I think the defendant is guilty of rape, but not of murder...'*

Mock Judge: *'Why?'*

J5 I2: *'I considered the wine bottle where there is DNA from both the defendant and the victim... The idea I had is that, eventually, the victim somehow accepted... I mean, it's possible they spent the night together...'*

Mock Judge: *'So you want to say that the intercourse was initially consensual...'*

J5 I2: *'No, no...'*

Mock Judge: *'So, technically...'*

J5 I2: *'No. I wanted to say: it's possible they spent the night together, without the victim's intention to have sex [...]'*

Although the already highlighted benefits of having a judge help jurors understand how to correctly reach a verdict cannot be disregarded, it has to be considered that potential side-effects of discouraging lay jurors to freely express their opinions may well derive from judges' tendency to interrupt. It may be the case that the judge's interruption, even if arising from a misunderstanding, intimidates the juror, who might end up not fully expressing their thoughts, if in contrast with the judge's; it may also happen that the misunderstanding created by the interruption is not promptly clarified and generates further confusion in the overall discussion. These and other potential risks have to be considered when evaluating benefits and limitations of the presence on the panel of a professional, who, perhaps driven by the awareness of their greater competence, might tend to interfere with the natural unfolding of the discussion even when not needed.

– Influencing remarks

The mock judge tended to express her opinions and to respond to jurors' claims in a quite assertive way, which, combined with her own and jurors' awareness of her role, often resulted in her being overly influential. For example:

Mock Judge (in I3): *'I'm sorry, I'll have to say this again, the signs of violence found on the victim's body are compatible with a violent sexual intercourse. This does not mean they are incontrovertibly indicative of the occurrence of a sexual assault. You must be critical in the analysis of the overall evidentiary framework.'*

Moreover:

J5 I2: *'Well, I still think there was rape...'*

Mock judge: *'Hmm... Do you? Are you sure? Beyond any reasonable doubt?'*

J5 I2: *'No, I think I'll acquit then.'*

Regardless of the content and intrinsic meaning of such mock judge's remarks (which are, not surprisingly, sensible), it is clear from the examples above that the tone and demeanour through which the messages are conveyed have an unavoidable impact on jurors' reactions. This emerges quite evidently from the last quote, which shows a sudden change of mind on the part of the juror, which did not occur after juror's further reflection (J5 I2 did not take enough time to do so), but in reaction to the mock judge's (rhetorical) questions. While it is beneficial that jurors are led to reflect on their opinions and are reminded of the legal standards by which they should abide (e.g. reasonable doubt), doubts arise when that further reflection is not prompted and jurors' changes of mind occur as a result of influential remarks put forward by the professional. This does not appear to be in line with the nature of the jury trial and should, therefore, be avoided.

– (Too) Early opinion disclosure

The mock judge's opinions, according to the Italian Code of Criminal Procedure and the interviewed judges' suggestions, needed to be expressed only after the lay members had space to express theirs. While this rule was observed in the context of voting

procedures, it was more difficult (at times impossible) to avoid that the mock judge's thoughts emerged during the natural unfolding of the discussion. In fact, despite the attempted control of this aspect, mock judge's opinions were perceivable from simple statements or answers to lay members' remarks, even at the very early stages of deliberation.

Indeed, it is difficult to imagine how a judge, especially when not purposely instructed and trained to exert a mere directive and informative role, would lead a discussion without letting their opinions emerge whatsoever. With regard to this disadvantage created by the judge's presence, it has to be concluded that, given the current rules, there was no way (in experimental settings as well as in reality) to prevent that from happening and to avoid the consequent influence exerted. This detrimental effect should be kept under control by providing the professional who sits on the panel with specific instructions in this respect.

– Control over topics addressed

The mock judge maintained control over the topics that were addressed, as that was part of her directive role. Also in this case, there were benefits (for example, her ability to focus attention on the relevant count of indictment when necessary):

Mock Judge (in I1): *'Ok, so far we've been focusing quite a bit on the rape accusation. We need to discuss the murder as well.'*

However, when considering the amount of overall control with which this provided her, risks of manipulation cannot be excluded. The mock judge quite often strategically used her leading role and consequent opportunity to move from one topic to another, in order to divert the conversation when it was not moving towards what she believed to be the right conclusion. Also regarding this behaviour, specific instructions might avoid, or at least, reduce unwanted influencing dynamics.

In the exercise of the difficult duty of directing a group of lay decision-makers in the criminal justice system, a judge should be able to find a balance between guiding the discussion according to the legal standards and leaving space for the contrasting

opinions of the lay members (Zambuto, 2016). Given the lack of instructions and specific training, judges tend, perhaps unconsciously, to overpower lay jurors. This comes as a result of the fact that the lay jurors face difficulties in understanding legal language, procedures and technicalities and in acquiring familiarity with an overwhelmingly complex environment. In such contexts judges, conscious of their own greater experience, easily end up influencing the majority of lay members (Lanza, 1994). It could be argued that, given the profound distrust and widespread criticism against lay juries, the judges influence might be a price that criminal justice systems should be willing to pay in order to grant defendants a fair trial. On deeper reflection, however, depriving the lay jurors of their decisional power would mean depriving the jury trial of its actual democratic component, and therefore of its most essential function.

4.4.2 Mock Juries' Deliberations – Quantitative/Statistical Analysis

4.4.2.1 Sample

Gender, Age and Nationality

The overall sample (excluding the mock judge) was composed of 34 females (67%) and 17 males (33%), distributed across four age ranges: 23 participants (45%) were in the ages range “25 or under”, 20 participants (39%) were in the age range “26-40”, 3 participants (6%) were in the age range “41-55”, 5 participants (10%) were in the age range “56 or older”.

Participants were of several different nationalities. However, in categorising the variables into SPSS, only two values were used to differentiate between “British and other” and “Italian”. For clarity, this was done because participants’ nationality did not form the basis of any analysis in this thesis, and the only relevant grouping condition was “British” and “Italian”, as those two categories reflected the two different experimental conditions. The specification “and other” for the first category was added merely because participants who deliberated in the British condition were from different nationalities (which was expected as they were mostly students recruited

within the multicultural environment of the University of Leicester), while participants in the Italian condition were all Italian.

Occupation and Jury Service

Similarly to nationality, a categorization of participants' occupation was made taking into account the specific information needed. In particular, for the purposes of grouping descriptive data, two categories were created: "Students" and "Non-Students". Accordingly, 26 (51%) were students and 25 (49%) were non-students. These data reflect the previously mentioned recruitment strategy whereby students were easier to recruit in England, where the researcher worked within the academic environment, whilst members of the general public were easier to recruit in Italy. The characteristics of this sample were considered and limitations acknowledged, as will be discussed in Chapter 5 (section 5.7).

Two other relevant characteristics were considered and are more widely discussed in section 5.7, where strengths and limitations of this thesis are addressed: the majority of the student participants were Criminology students, which, despite not being ideal as an experimental condition for this type of study, indeed unearthed very interesting points. Moreover, there were mock jurors (3 in the British condition, and 2 in the Italian condition) whose occupation was related to the legal profession (trained barristers, forensic science experts, etc.), which was also an interesting aspect to take into account when considering to what extent these people (jury eligible members also in real life) were able to make reasonable, well-informed verdict decisions as well as making sensible comments during the course of the discussions and/or in their responses to questionnaires.

Similarly, it was recorded whether participants had already sat on a (real) jury in the past, in order to see whether that would generate relevant differences in their approach and attitude. Only three participants (all in the British condition) stated that they had served on a real jury, which was too small a number to use in further statistical analysis.

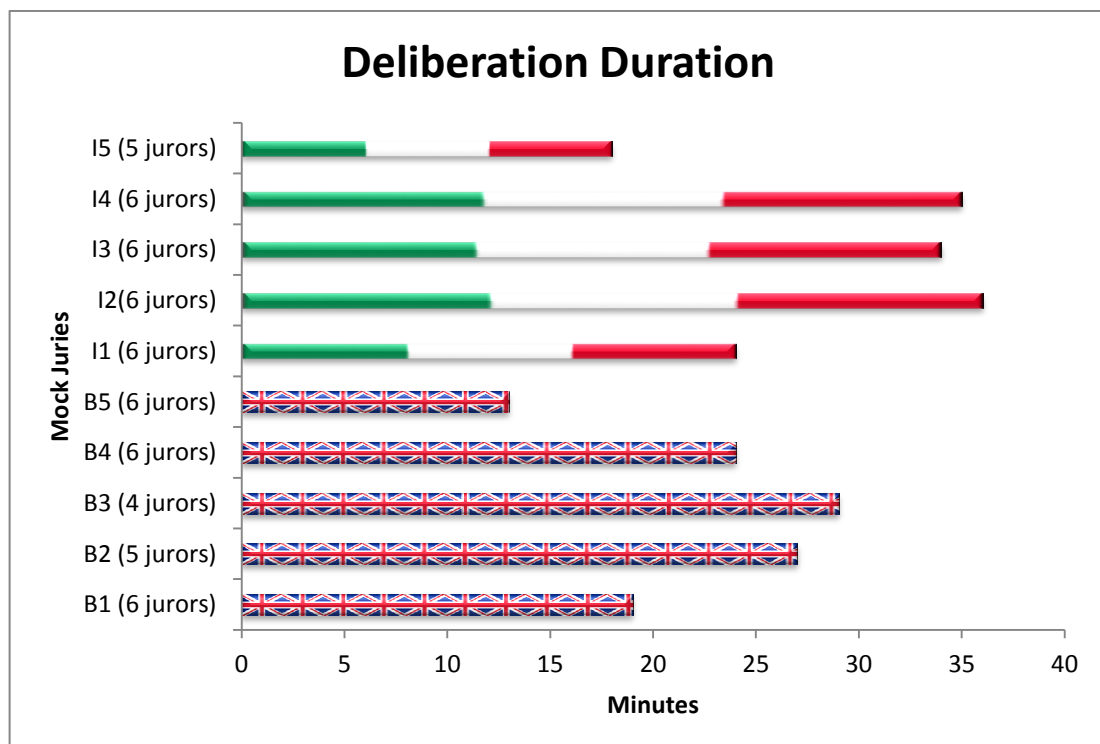
British condition and Italian condition (composition)

The overall sample consisted of 27 participants in the British condition (53%), and 24 participants in the Italian condition (47%). Given the equal number of groups (five) per condition, a corresponding equal number of participants was expected, however, the slight difference reported is due to two factors: firstly, a few participants not attending; secondly, the presence of a mock judge in the Italian condition. Specifically, three participants did not attend British sessions (one in B2 and two in B3), and one participant did not attend an Italian session (I10). The presence of the mock judge also affected the numbers because that participant was not included in the analyses; therefore, even though each Italian group was formed of six members, only five per group resulted for the purpose of the quantitative data collection and analysis.

4.4.2.2 Deliberation duration

All mock juries were told they could deliberate for approximately 30 minutes, with the further specification that they should freely discuss the issues at hand and it would have been the researcher's duty to interrupt the discussion if needed. Deliberations' length was recorded from the videos:

Figure 4.1



As Figure 4.1 shows, Italian mock juries' deliberations lasted slightly longer: the average duration for deliberations in the Italian condition was 29 minutes, while the average duration for deliberations in the British condition was 23 minutes. The result is not surprising, when considering that the presence of the mock judge in those cases gave rise to a series of deliberation activities that were not carried out in the British condition (e.g. reading of the summary report, explanation of the legal principles). Overall, only in three instances (I2, I3, I4) the researcher's intervention was needed to ensure, for the sake of experimental consistency, that those discussions did not last much longer than the others. The approximately set deliberation time was indicated to guarantee that the overall activity's length (case description and questionnaires completion included) was consistent with the informed consent form indications (Appendix C), and without neglecting the unrealistic nature of such (required) deliberation time restriction (Ellison and Munro, 2010a). The choice of a relatively short deliberation time was mainly made in light of the need to avoid discouraging participation: for the purpose of the experiments, having deliberations, albeit not long, was better than having none for lack of participants. Nonetheless, the potential impact of this element on the findings was considered when designing the experiments and in the analysis phase. In designing the experiments the amount of time was, therefore, determined through a balanced assessment of what would be an acceptable duration for potential participants and what would be the actual time required for the (experimental) discussions given the nature of the experimental stimuli. In this respect, it was considered that the, elsewhere highlighted, ambiguous nature of the fictional crime case scenario along with the weakness of the evidence presented would have left room for discussion, but also led to a point where no more discussion would be needed, once that weakness had emerged as a sensible reason to acquit. Indeed, the fact that in the majority of cases, deliberations ended naturally suggests that, for the proposed experimental stimuli, that amount of time was appropriate to reach a verdict.

4.4.2.3 Verdict preferences

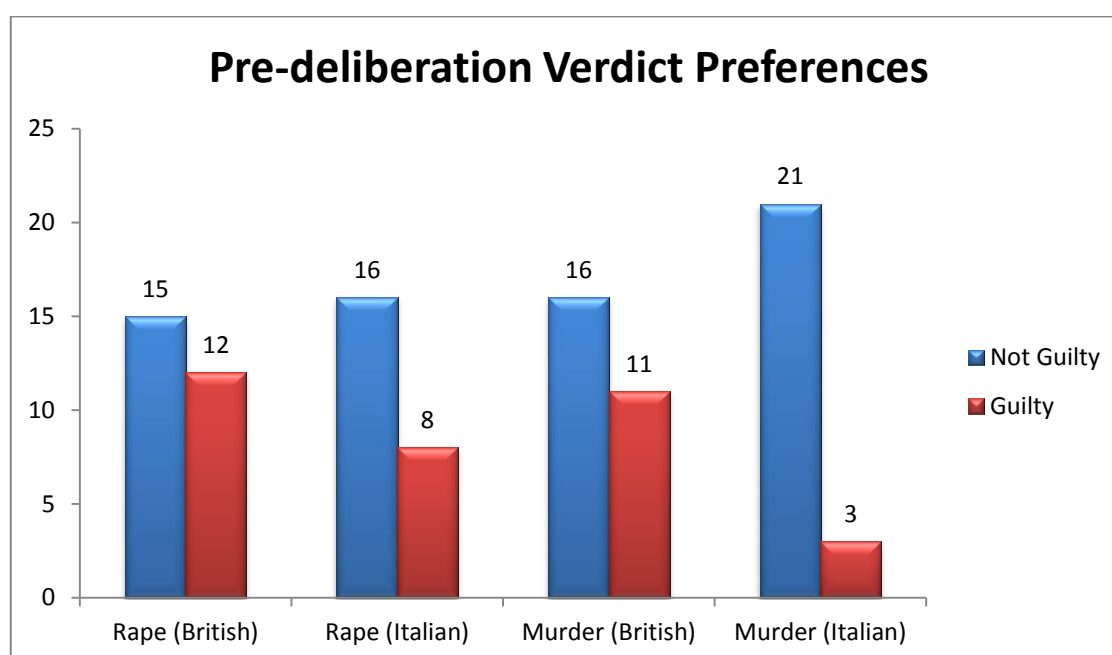
All groups across the two conditions eventually reached a verdict of not guilty on both counts (murder and rape). This result was foreseen as a plausible outcome and,

accordingly, can be explained. This occurred because, for experimental purposes, the hypothetical criminal case scenario used was intentionally designed to be ambiguous and to lack sufficient evidence for conviction. The inconsistency and the errors in the evidence were so significant (no real case would probably present such errors) that a “not guilty” verdict was the only reasonable conclusion to reach.

The fact that mock juries used for the experiments reached the same verdict is neither new nor peculiar to the experiments presented in this thesis. Other studies have had the same result (e.g. Charron and Woodhams, 2010) and the fact that, notwithstanding this, there still was room for discussion (in past studies as well as in the present experiments), can be justified in light of Ellison and Munro’s (2010a, p.75) remark that ‘juries, when faced with the same trial stimulus, can not only reach divergent verdicts, but can embark upon radically different routes to reach the same destination’. The different routes upon which mock juries in these experiments embarked constituted the main elements of the present investigation, as it has been previously reported through the qualitative analysis in this chapter and as will be further done here through quantitative analysis.

Besides the fact that heated discussions happened, further proof that the hypothetical case used was an effective experimental tool is given by the fact that no mock jury in the experiments started deliberation from a unanimous verdict. There was indeed a broad spread of pre-deliberation opinions, which were recorded and collected through the pre-deliberation questionnaires (and, when expressed, double-checked in the videos). Descriptive statistics revealed that, 31 participants (61%) across both conditions started with a pre-deliberation verdict of “not guilty” for rape, while 20 participants (39%) considered the defendant guilty of rape. Regarding the accusation of murder, 37 mock jurors (73%) started from a “not guilty” verdict, and the remaining 14 (27%) considered the defendant guilty. This confirms that mock jurors, individually asked about their opinions pre-deliberation, were not unanimous in their verdict preferences. Figure 4.2 shows the distribution of pre-deliberation verdict preferences by condition for each count of indictment.

Figure 4.2



Italian mock jurors were slightly more lenient from the outset. This is particularly evident with regard to the murder accusation, where only 3 mock jurors out of 24 believed the defendant had killed the victim. Given that this verdict preference was taken before deliberation started, it is implausible to claim that the presence of the judge or any other experimental variable might have had an impact on this and could justify these differences. All the elements used before the beginning of deliberation were consistent across conditions and therefore no definite conclusion can be drawn regarding the greater leniency shown by Italian judges in their pre-deliberation verdict preferences. Furthermore, besides the pre-deliberation verdict preferences, individual verdict preferences were recorded also post-deliberation and they allowed conclusions to be drawn in terms of conformity dynamics across groups and conditions.

4.4.2.4 Conformity

Since it is an intrinsic characteristic of groups' interactions and due to the specific rules set for the completion of the deliberation task in this study (e.g., unanimity was required), it is easily understandable that some degree of conformity – whether private or public – was indeed expected to occur in the mock jury experiments. That mock jurors conformed was confirmed by the data, when considering that all the

groups reached a unanimous verdict, but none started from a unanimous position. In order to understand whether individual mock jurors privately or publicly conformed, pre- and post-deliberation questionnaires were used to deduce conformity type. This was achieved by asking mock jurors, in both questionnaires (Appendix I), what their individual opinion was (regardless of the group verdict, in the post-deliberation questionnaire). It was therefore possible to deduce when a change of mind had occurred, and whether it was best described as private or public conformity.

In order to analyse this aspect, four variables were initially created to indicate mock jurors' pre- and post-deliberation verdict preference on each count (guilty or not guilty, of rape and murder). A further variable was created to determine any change of mind which occurred and in what "direction" it went – that is, from not guilty (NG) to guilty (G) or *vice versa*, as well as no change of opinion. Findings at this stage demonstrated that the majority of mock jurors (86%) did not change their mind. However, this was not sufficient to know what type of "no change of opinion" that was in each case. Therefore, it was necessary to take into account a further difference between mock jurors who did not change their mind, because their choice corresponded to the final verdict from the outset, and jurors who did not change their mind, because – although formally agreeing with the final verdict as they all did – they privately kept their verdict preference. Accordingly, a further analytical step was taken, which accounted for this difference and led to the creation of three categories: one including mock jurors that did not need to conform (did not change opinion, because they chose not guilty pre- and post-deliberation); one including mock jurors who privately conform (changed opinion privately, as reported in post-deliberation questionnaire); and one including mock jurors who publicly conformed (did not change their opinion privately, as reported in post-deliberation questionnaire). This was observed separately for each of the two counts of indictment.

Figure 4.3

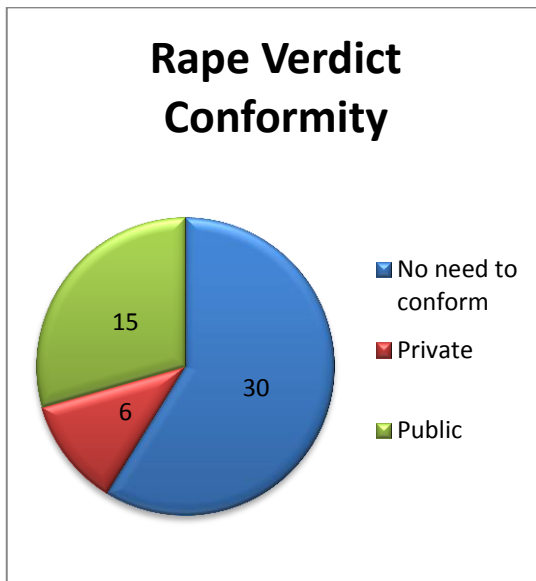
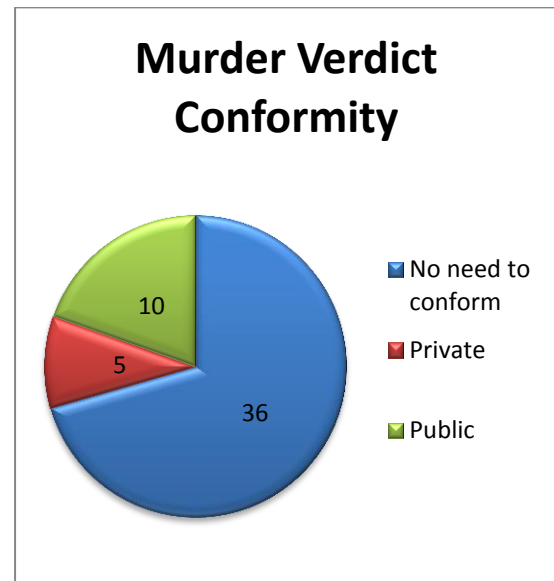


Figure 4.4



Figures 4.3 and 4.4 show that the majority of mock jurors did not need to conform, while the majority of those who conformed, did it only publicly. Furthermore, it was considered relevant, for the purpose of the present analysis, to also understand these dynamics within conditions, to see whether any difference could be attributed to the variables under analysis.

Figure 4.5

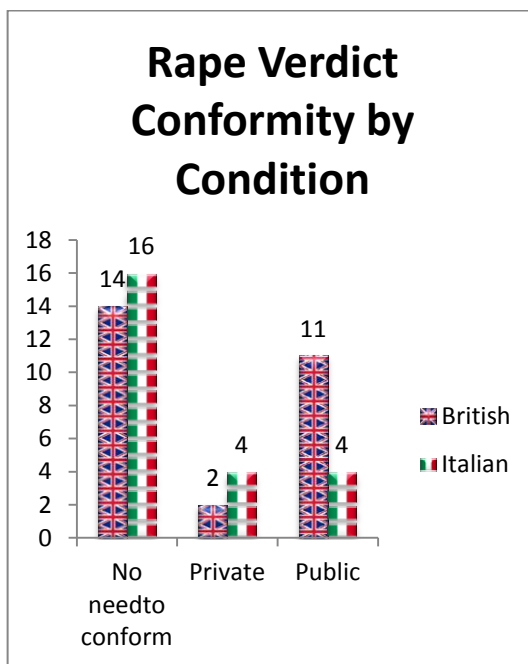
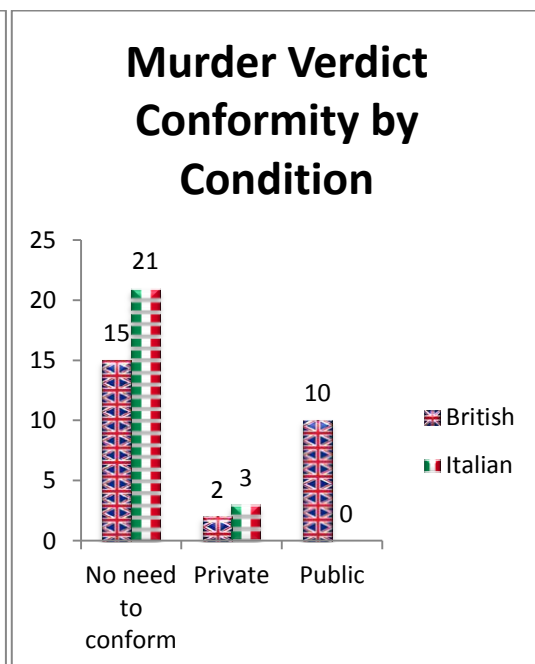


Figure 4.6



The descriptive analysis conducted to more closely observe conformity dynamics in mock juries in the two conditions revealed interesting results. As Figures 4.5 and 4.6 show, there was no remarkable difference between the tendency to conform on either of the counts of indictment. For both rape and murder verdicts, the majority of people in the two conditions started from a not guilty verdict, which, given the very weak and ambiguous nature of the evidence presented, is perhaps not surprising. Moreover, almost equal numbers of mock jurors privately and publicly conformed for both rape and murder innocence verdict. In general, the higher leniency of Italian jurors, already registered in pre-deliberation choices appeared confirmed by this further analysis (as shown by the “No need to conform” column).

Although the proportion of participants who conformed was not high enough to run any inferential statistics, interpretative remarks can still be proposed to provide explanations for the results. First of all, it cannot be disregarded that, across the two conditions and for both counts of indictment, private conformity occurred less often than public conformity. This confirms the basic assumption of conformity tendency, which is that people are generally willing to comply with the most desirable opinion, which in the context of jury decision-making, should be the most reasonable and fair, from a legal point of view. However, people are far less willing to modify their private opinion, which seems to confirm what previous research has claimed: jurors tend to have their minds made up from the beginning (since the first ballot and, at times, even before all evidence is presented), and to maintain these views until the end of deliberation (Kelman, 1958; Kessler, 1975).

Furthermore, mock jurors in the Italian condition tended to conform privately more often than mock jurors in the British condition, to the extent that, for the murder accusation, all the Italian mock jurors who changed their opinion during deliberation, did so by privately changing their mind. In contrast, the British mock jurors who conformed did so mainly publicly. In this respect, presence of the mock judge and requirement for verdicts’ motivations in the Italian condition may have had an impact and, therefore, further reflections on this point will be proposed in the final discussion

(Chapter 5) of this thesis in the context of a general consideration of the impact of the two variables on jury decision-making.

4.4.2.5 Confidence

In comparing British and Italian mock jurors' verdict preferences, another aspect that was measured pre- and post-deliberation was their confidence with their choice. Mock jurors were asked, in both their pre- and post-deliberation questionnaires, to rate their confidence with the verdict preference expressed on a 5-point Likert scale (Appendix I). This variable was measured in order to explore multiple aspects of mock jurors' confidence, namely, whether and to what extent their mean confidence varied before and after deliberation across the conditions. It was hypothesised that the discussion with other members of the panel could affect the individual levels of post-deliberation confidence, such that the mean confidence could be increased or decreased after deliberation. Similarly, it was hypothesised that the presence of a mock judge and the requirement for motivated verdicts may generate a further variation in post-deliberation confidence across conditions.

Two Independent-Samples T-Tests were performed to determine whether levels of confidence were affected by deliberations and/or by the impact of the two independent variables manipulated in the two conditions. Contrary to the hypotheses, the results showed that confidence variations, both overall and across conditions, were very small and not statistically significant:

Table 4.2: Confidence pre-/post-deliberation across conditions

	BRITISH	ITALIAN	TOTAL	<i>t</i>	<i>df</i>	<i>p</i>
PRE-DELIBERATION CONFIDENCE	3.26 Range values (min/max): 2-5	3.04 Range values (min/max): 1-4	3.16	1.062	49	.294
POST- DELIBERATION CONFIDENCE	3.56 Range values (min/max): 2-5	3.88 Range values (min/max): 2-5	3.71	-1.585	49	.119

According to the findings, it has to be concluded that neither deliberations per se nor presence of a judge and requirement for motivated verdicts had a significant effect on jurors' confidence. However, given that the analysed variables exerted an impact on several other aspects measured in this study, it has to be concluded that this lack of significant impact is strictly related to jurors' confidence. On reflection, this is not difficult to understand, when considering the difficulty of the task that jurors undertake. The lack of familiarity with the legal environment and rules, the objective uncertainty of the to-be-judged situation, the burden of making a decision where somebody else's life is at stake are all factors that cannot be easily overcome. Presumably, these and other elements of difficulty remain embedded in the decision-making process and act upon jurors' confidence generating the result that, even after discussing with the group, or receiving the help of a professional, or being driven towards deeper reflection in order to provide a legally-oriented motivation for the verdict preference, individual jurors are not very confident with their choice.

4.4.2.6 Perception of verdict fairness

In the post-deliberation questionnaire, mock jurors were also asked to provide their opinion on the fairness of the group verdict. The question, offering the participants a

further opportunity to express an individual opinion that they might have not felt comfortable enough to voice publicly, represented a confirmation of their personal verdict preference. Clearly, mock jurors who did not privately agree with the final verdict, would likely have found it to some extent unfair. Mock jurors were required to rate the fairness of the final decision reached by the group on a 5-point Likert scale (Appendix I). An Independent-Samples T-Test was conducted to determine whether opinions from mock jurors in the two conditions differed significantly. Results indicated that there was a significant difference between the scores for British ($M = 3.52$, $SD = 1.18$) and Italian ($M = 4.13$, $SD = 0.53$) conditions, with $t\text{-value} = -2.298$; $df = 49$; $p\text{-value} = .026$ ($p < 0.05$). According to these findings, Italian mock jurors found the group verdicts fairer on average than the British jurors did.

Consistent with what has already been observed regarding the greater tendency of Italian mock jurors to privately conform to the verdict decision, and as a further confirmation of the explanation provided for that result, the fact that those mock jurors were also more likely than the ones in the British condition to consider the verdict decision fair has to be attributed to the presence of the mock judge and/or the requirement for motivated verdicts. With specific regard to verdict fairness, on multiple occasions the mock judge focused the discussion on legal principles and explained that the group was not required to reach certainty about the defendant guilt or innocence, but rather to find enough evidence-based reasons to decide on one verdict or another. Plausibly, British mock jurors, who did not receive any explanation, continued to believe that, if they intimately thought the defendant was guilty while the group found him not guilty, that final decision was perceived to be less fair. Instead, a repeated explanation of the principle "*in dubio pro reo*" [Latin for: "when in doubt, (decide) for the accused"], and of its more profound foundations within a criminal justice system that needs to abide by certain standards, certainly conveyed effectively to Italian mock jurors the concept that a fair judicial decision is not necessarily the one that best reflects the opinion of the individual jurors, but it is the one that is taken in deference to the system's fundamental rules.

4.4.2.7 Leadership Perception

In addition to observing leadership dynamics through the video-taped deliberations and providing a qualitatively analysed account of them (section 4.4.1.5), leadership was investigated through how it was perceived by the mock jurors who experienced the deliberations. This, through the benefits of triangulation, enabled a determination of whether mock jurors' perceptions of leadership corresponded to what actually happened.

All mock jurors were asked in their post-deliberation questionnaire: "Were there members of the panel who tended to lead or prevail in the discussion?". This categorical (Yes/No) question was intended to gather information about all mock jurors' perception of leadership dynamics and to detect potential differences between the two conditions. The responses to this question in the two conditions were completely opposite. In the British condition, 89% of participants declared they felt the presence of a leader in their group, while 89% of participants in the Italian condition declared they did not perceive the presence of a leader in their group. A Chi-square test was run to determine whether or not this difference in the perception of leadership observed by condition were statistically significant. The association was statistically significant, with $\chi^2 = 36.49$ ($p < 0.05$).

Given the different composition of the two types of panels and the presence, on the Italian panels of the (mock) judge, who is considered a predetermined leader (Zambuto, 2016), the result reported above might appear, at first glance, contradictory: the groups with an obvious, predetermined leader claimed they did not perceive a leader to be present, while the groups composed of peers, where everyone should have been on the same level, widely felt leadership dynamics to be in action. Nevertheless, at closer inspection, these findings may reasonably be explained, considering that, as will be clarified, the conditions in which the groups operated may well be the very reason for this difference.

Preliminarily, even before focusing on the two specific conditions, it has to be considered that the mock jurors personal perception of what happened in their

deliberation room might not reflect accurately what happened in reality. In an effort to investigate this aspect, post-deliberation questionnaire responses were cross-checked against the results of the observation conducted about leadership dynamics (see section 4.4.1.1.4), and in the majority of cases the perceptions did not match the reality:

Table 4.3: Observed vs. Perceived Leadership in British condition

MOCK JURIES	QUALITATIVE ANALYSIS: OBSERVED LEADERSHIP DYNAMICS	QUANTITATIVE ANALYSIS: "WERE THERE MEMBERS OF THE PANEL WHO TENDED TO LEAD OR PREVAIL IN THE DISCUSSION?"
B1	No leading tendencies	5 YES 1 NO
B2	2 contrasting leading personalities	4 YES 1 NO
B3	No leading tendencies	3 YES 1 NO
B4	2 "quiet" leading personalities	ALL YES
B5	1 "strong" leading personality	ALL YES

Clearly, some of the perceptions do not match the reality of the observed deliberations. This, however, can be interpreted in light of several elements. Despite any effort made to make the observational process as objective and systematic as possible (as described in section 4.4.1.1.4), a certain degree of researcher subjectivity cannot be excluded from any interpretative analysis (Drapeau, 2002), and thus some of the discrepancies might be attributed to this factor. However, given the conclusions drawn by the literature regarding people's leadership perception, explanation can be plausibly found elsewhere. People's implicit theories of leadership (Eden and Leviatan, 1975) challenged the longstanding assumption that individuals, in order to identify and

evaluate their leaders, relied on actually observed behaviours. By contrast people's implicit theories of leadership constitute their preconceptions about leaders' traits, behaviours, and consequences of their behaviours (Emrich, 1999). Eden and Leviatan's (1975, p.741) findings led them to conclude that 'leadership factors are in the mind of the respondent'.

Considering these findings, and even considering that perception itself is an intimate process that may be based on factors that are not perceivable externally, doubts arise about the fact that mock jurors' leadership perception could actually be trusted to represent what happened in reality. This, by extension, could apply to the Italian mock jurors as well, who, even in the presence of an officially defined and predetermined leader, declared that they did not perceive dynamics of leadership to be unfolding during the discussions. However, given the recorded statistical significance of the association between leadership perception and condition, this explanation cannot be accepted as satisfactory as it would not account for the above-mentioned results. Nevertheless, an alternative explanation can be found when reflecting on the role of the jury judge and its implications.

As observed elsewhere in this thesis (Chapter 3), the role of the judges who work with lay jurors is – at least theoretically – a directive role. That is what makes those judges predetermined leaders. Within the directive role of these leaders (as the interviewed judges claimed – see Chapter 3) falls also their duty to control the balanced unfolding of the discussion, aimed at guaranteeing that everyone has enough space to freely express their opinions. The presence of a judge, thus, ensures that potential conflicts are solved through an open discussion, and, when necessary, confrontations are placated due to his/her authoritative intervention (Berti, 2002). It is then possible that judges, who are not seen as an integral part of the group as their authority is already conferred to them by the system (Zambuto, 2016), end up not being perceived as leaders, but rather as “external” members. Accordingly, the mock judge in the experiments was not seen as a leader, and yet, by simply exerting her directive role, did not allow anyone else to prevail, hence mock jurors did not perceive anyone as a leader.

Albeit sensible, this conclusion automatically implies that the Italian mock jurors were not referring to, and indeed were not even considering, the mock judge when responding to the leadership question. In fact, due to the way the question was posed (perhaps, a methodological flaw that could be improved in future empirical experiments), it was not possible to know whether, in answering that question, participants were referring to the mock judge or any other lay participant. Notwithstanding this issue, the remarks made above regarding the peculiar position of the judges and its implications in terms of leadership, along with the observed tendency of the mock judge in the experiments to involve in the discussion the more silent members (which presumably contributed to Italian jurors perception of lack of leading tendencies, as they were prevented from happening), seem to suggest that the solution proposed above is still the most reasonable explanation for the reported statistical findings.

4.4.2.8 Need for/usefulness of a judge or someone to direct the discussion

To continue assessing mock jurors' perception of deliberation factors and dynamics, two more questions were asked about the need for and usefulness of the presence of a judge and/or of someone to direct the discussion. Given the fundamental difference between the two experimental conditions in this respect (Italian mock jurors had a mock judge on the panel, whereas British jurors did not), the two questions had to be asked in a different, yet corresponding, way (as reported below). Moreover, the distinction made between a "judge" and "someone to direct the discussion" was also purposely planned to investigate slightly different aspects. Indeed, while in the case of Italian juries, the two functions referred to the same person, considering that the presiding judge has also a directive role, in British juries, the lack of such a figure leaves space to imagine that the two roles do not need to necessarily correspond.

Figure 4.7

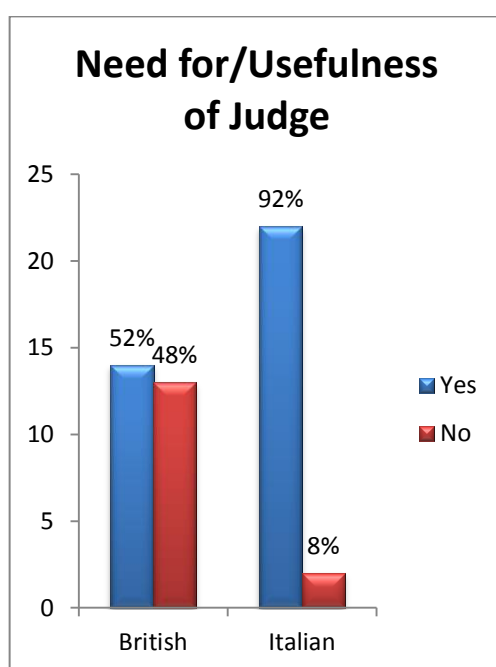
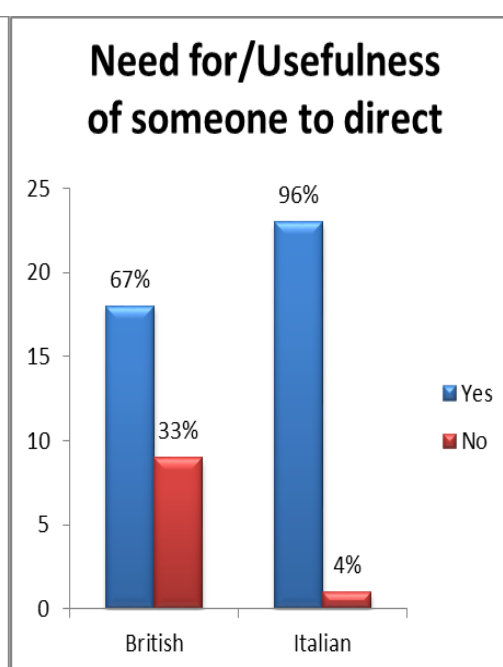


Figure 4.8



With regard to Figure 4.7, British mock jurors answered the question: ‘Do you think the presence of professionals (i.e. judges) on the jury panel might have been of help for you?’, while Italian mock jurors answered the corresponding question: ‘Do you think that the presence of a judge on the panel has been of help for you?’. With regard to Figure 4.8, British mock jurors were asked: ‘During the deliberation, did you think there was need for someone to direct the discussion?’; Italian mock jurors were asked: ‘During the deliberation, did you find it useful to have someone on the panel to direct the discussion?’.

Chi-square tests were performed and showed that, in both cases, the observed differences between conditions were statistically significant, with $\chi^2 = 9.702$ ($p < 0.05$) in regards to the findings in Figure 4.7; and $\chi^2 = 6.857$ ($p < 0.05$) in regards to the findings in Figure 4.8. This confirmed that Italian mock jurors scored the usefulness of the presence of either a judge or someone to direct the discussion as more positive compared to British jurors. Despite the higher percentages of “Yes” rather than “No” answers to both questions also within the British condition (52% and 67%), British mock jurors expressed the need for either a judge or someone to direct the discussion lesser than Italian mock jurors, when percentages are compared across conditions.

These findings suggest that those mock jurors who did not have a mock judge on the panel felt less need to have one and/or at least someone to help them by directing the discussion, yet those mock jurors who did have a judge on their panel seemed to show greater appreciation of their helpfulness. This could be reasonably explained by the fact that, seeing the judge in action, those mock jurors had a concrete idea of how important that presence and directive role was. Plausibly, mock jurors who are used to the idea of purely lay juries do not know what participation of a judge might consist of, whereas mock jurors who benefitted from that help, had a clear idea on which to base their opinion.

Even though British mock jurors appeared to appreciate less than Italian mock jurors the usefulness of either a judge or someone to direct the discussion, overall they (as well as the Italians) realised that having someone to help them would be better than deciding on their own. This might result as a natural consequence of the difficulty of jurors' task. It should not be surprising in fact, given the widespread criticism against the lack of competence and of legal knowledge on the part of jurors (Reifman, Gusick and Ellsworth, 1992), that they would feel better if they had someone who could direct them and/or provide them with the further information needed. This may be a way to make better informed decisions, or to partially shift responsibility, but regardless of the reasons, it is understandable that jurors perceive help as a benefit in facing their daunting task.

4.4.2.9 *Motivation*

The other fundamental element of procedural difference that this thesis investigated was the presence/absence of the requirement for motivated verdicts. As previously described (section 4.2.4.2), to reflect this difference in the experiments, Italian mock jurors were explicitly instructed, since the outset of deliberation, that they had to provide motivated verdicts (and the requirement was also constantly reiterated by the mock judge throughout deliberations). No similar specification was made to the British mock jurors, who, instead, were left free to deliberate as purely lay juries normally do. However, for the purposes of the comparative investigation, all jurors in both conditions had to answer a question in their post-deliberation questionnaire about the reasons underlying their verdict preference (Appendix I). This provided an opportunity

to understand the nature of the reasons that supported mock jurors choices, and to test for differences across conditions in this regard.

All mock jurors provided brief open-ended answers, which in the majority of cases consisted of blunt statements (the most common was, for example, '*Not enough evidence to convict*'), which did not provide the richness required for qualitative analysis. Consequently, with the aim to use it for quantitative analysis, the responses were all coded into the variable "Legally-oriented motivations", so that, depending on the content of their answers, mock jurors motivations were classified as based on relevant legal reasons or not. To further specify, responses such as: '*Not guilty due to lack of evidence*' (J3 B1) was classified as legally-oriented, whilst responses such as: '*Piggott looks to have done something unpleasant [...]*' (J2 B4) were classified as non-legally-oriented, in that they were based on legally irrelevant observations. Crosstabulations yielded the following results.

Figure 4.9

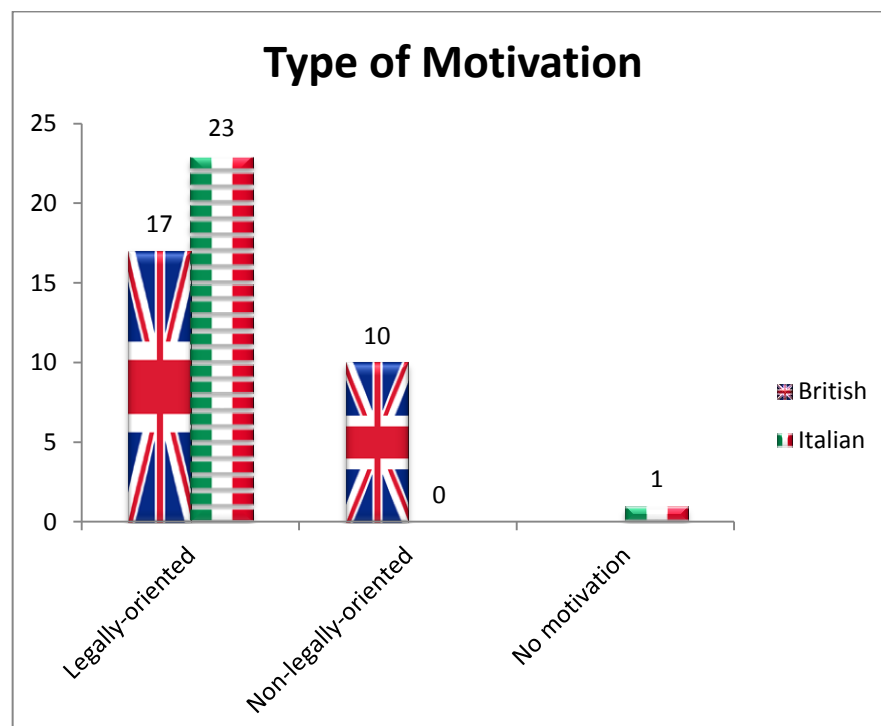


Figure 4.9 provides an account of the type of motivations provided by mock jurors in the two conditions. Despite the lack of any specification about verdict preference

motivations and notwithstanding the absence of a mock judge, 63% of the British mock jurors provided legally-oriented motivations. This result seems to suggest that, even without being prompted to think about the reasons underlying their verdict preference and without receiving any guidance during the deliberations, the majority of British mock jurors were able to articulate their decisions based on legally relevant values. Whilst this may certainly be possible, and there are those who would argue that jurors are perfectly capable of doing their job properly (Grove, 2012), some circumstances connected to the experiments conducted for this thesis should suggest caution in drawing definite conclusions in that sense.

One of the issues that this study presented is that some participants were Criminology students. Thus, it is possible that those participants' greater familiarity with legal matters may have increased their tendency to look for legally-oriented motivations. On the other hand, it would be probably incorrect to conclude in the opposite way, that is that these participants performed too well to be considered lay mock jurors, because, as evidently emerged by the analysis conducted so far, their reasoning and competence appeared flawed and lacking rigour in most occasions.

Far less ambiguous appear to be the findings regarding the nature of motivations indicated by the Italian mock jurors, who, excluding one case of lack of motivation (J2 I2), all provided legally-oriented motivations. Indeed the difference between the conditions was statistically significant with $\chi^2 = 11.764$ ($p < 0.05$). Mock jurors in the Italian condition, even when they publicly conformed to reach a unanimous verdict (and, therefore, kept their original idea), included in their motivations reference to the legal principles in support of their decision to acquit. It is, consequently, correct to conclude that both the requirement for motivated verdicts and the presence of the mock judge had an impact on these mock jurors' reasoning, leading them to better understand how to make a legally fair decision. Consistent with literature on explanation-based decision-making, mock jurors who knew they had to externally express their reasons (during deliberation and in the questionnaire) tended towards the choice that would be the most justifiable to other members of the group and by superiors (Simonson, 1989), which in the context of these juries was the most legally

appropriate choice. This has been accompanied, throughout deliberation, by the constant intervention of the mock judge, who by asking jurors to express their motivating reasons and by inviting them to make those reasons legally relevant, has plausibly generated some degree of self-awareness within mock jurors, stimulating their need to find legally-acceptable reasons.

4.5 Conclusion

In order to understand and assess the potential impact of the presence of judges and requirement for motivated verdicts on jury decision-making, mock juries were assigned to two different experimental conditions: British mock juries deliberated according to the British jury trial rules (only lay people, rendering unmotivated verdicts), whereas Italian mock juries deliberated according to the Italian jury trial rules (mixed jury, rendering motivated verdicts). The audio-/video-recording of the mock juries' deliberations, analysed in conjunction with individual jurors' responses to pre-/post-deliberation questionnaires offered varied and multifaceted insights into the matter. The triangulated, mixed-method analysis conducted covered a series of areas of relevance to both individual and collective decision-making, with the aim to understand to what extent the two manipulated variables affected it. The comparison between the two conditions produced numerous differences (in both the qualitative and quantitative analysis), which suggest that indeed the variables influence jury deliberations. Considerations on the implications of the effects found are provided throughout this chapter and further reflections will be proposed in the context of a final overall discussion in Chapter 5.

CHAPTER 5

FINAL DISCUSSION

*'Consider what you think justice requires,
and decide accordingly.
But never give your reasons;
for your judgment will probably be right,
but your reasons will certainly be wrong.'*

Lord Mansfield

5.1 Introduction

The aim of this final chapter is to propose conclusive reflections on the findings of this research, in order to show how they provided answers to this thesis' overarching research question, that is whether and to what extent procedural differences of the British and Italian jury trial may affect jury decision-making during deliberations. Therefore, while previous chapters have provided and discussed answers to specific research questions and sub-questions, in this chapter, findings from the two studies conducted – Italian judges interviews (Chapter 3) and mock jury experiments (Chapter 4) – will be collectively considered in the context of an all-comprehensive discussion.

Through the employment of a mixed-method design and making use of the benefits of triangulation, findings yielded from both qualitative and quantitative data were reinforced and complemented each other. Results from the conducted studies confirmed the most widely recognised theoretical and empirical findings of previous literature on the topic. Moreover, the comparative approach adopted in this thesis provided new points of analysis and opportunities for reflection. Since the two studies investigated some of the same aspects from different perspectives and through different methods, their contributions will be merged in the overall discussion that follows.

The impact of the two independent variables on three broad areas – jurors’ behaviour, jurors’ errors and jury deliberation dynamics – will be discussed, as resulted from the findings. Following, further reflections will be made on the two variables individually considered. To conclude, mention of the contribution of this research to the wider literature is made, along with a comprehensive account of strengths and limitations that the research presents. Considerations made in this respect should be taken into account in future research endeavours, as advised in the context of the provided future research suggestions.

5.2 Impact of composition and motivated verdicts on jurors’ behaviour

Findings from both studies are considered in order to draw conclusions about the impact of the two independent variables on jurors’ behaviour. For the purpose of this discussion, this will entail jurors/juries approach to deliberation, narrative construction, confirmation bias. Study 1 and 2, to different extents and from different perspectives, have both offered interesting points of reflections on these aspects.

Regarding jurors/juries approach to deliberations, findings from the mock jury study (Chapter 4) showed that, when left free to decide how to begin deliberation, purely lay mock juries tended to opt for a “verdict-driven” or “evidence-driven” approach and that, moreover, regardless of the choice, at times the adopted approach resulted in a mix between the two. This is consistent with previous literature that rejected the clear-cut distinction between the two approaches (Ellison and Munro, 2010a). However, consistency was observed in mock juries with mixed composition (Italian condition), where the mock judge, following the set voting procedure (CPP, 2010) and instructions based on real judges accounts (see Chapter 3), asked each lay member to declare what their vote was and to explain their reasons.

The observable difference is most likely to be ascribed to the presence of the mock judge who directed the “go-round” procedure, in that there are no reasons to plausibly believe that, lacking that presence, Italian mock juries would have maintained that

consistency (since British lay mock juries did not). Regarding the potential impact of the requirement for motivation, the fact that the mock judge in the Italian condition started asking for motivations at this early stage of deliberation may well have “planted the seed” in the minds of the jurors and oriented them towards thinking about their motivating reasons rather than expressing their preferences on the basis of first impressions and other legally irrelevant factors. However, such an impact could not be measured at this stage of deliberation, given that discussions had not started yet, and therefore definite conclusions cannot be drawn in this sense.

Given the extensively proven impact that approaches to deliberations have on following deliberation discussions (Ellsworth, 1989; Ellison and Munro, 2010a), the benefits generated by the degree of consistency that the judge’s presence conferred are undeniable. However, based on previous research, early votes can have detrimental effects on deliberations (Whitworth and McQueen, 2003), leading jurors to focus on defending those votes rather than remaining open to sensible and at times necessary changes of mind (Hastie, Penrod and Pennington, 1983). Therefore, it is reasonable to conclude that, if a certain degree of consistency appears beneficial, this should perhaps move in the opposite direction, that is assuring that juries begin discussing the issues and vote at a later stage of deliberation.

Another risk connected to the “direction” that jurors’ reasoning follows was observed in their tendency to develop, at times, arguments starting from an (unsupported) conclusion they believed to be true and going backwards to find motivating reasons. This aspect was not specifically addressed in Study 1, but was apparent in the results of Study 2. The recorded phenomenon, identified also in the previous literature which emphasised how people reason bi-directionally (Glöckner and Engel, 2013; Simon, Snow and Read, 2004), was observed in the presented mock jury experiments in cases when a conclusion was initially presented and an explanation was found only afterwards. In such cases the “direction” of the reasoning was clearly not from the evidence to the conclusion (as it should), but also from the conclusion to the evidence, which affected the very perception and interpretation of evidence (Holyoak and Simon 1999; Thagard and Millgram 1995; Glöckner and Betsch, 2008). This happened

because, when jurors started from the conclusion to evaluate the evidence, their view of the evidence was biased such that they tended to assign more weight to the evidence that supported their conclusion (Nickerson, 1998; Glöckner and Engel, 2013).

The impact that the requirement for a motivation may have had is self-evident. Although such a requirement did not completely prevent the natural tendency towards confirmation bias from entering Italian mock jury deliberations, it can be argued that it reduced its occurrence by making the mock jurors aware of their accountability and by prompting them to focus on their motivating reasons, before reaching rushed conclusions. Less impactful was the role of the mock judge in these cases, because these dynamics may easily go undetected in a discussion, unless the later explanation is particularly imaginative and unrealistic. In such cases, the nature of the explanation reveals the faulty inverse reasoning, which the judge could then correct. The absence of requirement for motivation and/or of a judge creates, instead, a situation where such inverse reasoning may well go undetected for jurors who are neither held accountable nor to some extent guided.

Clearly, how jurors reason has an effect on deliberations in several ways and at several stages of the discussion. The considerable complexity of the cognitive effort required make it self-evident and has been at the heart of most research in the field and of this thesis as well. The most widely accepted and empirically tested model of jury decision-making, the story model, highlights jurors' tendencies to inject into the decision process their own prior knowledge, preconceptions and expectations. Jurors make sense of the case facts, assembling the evidence into a story that provides a coherent causal explanation which is conducive with their past experience, knowledge, etc. (Hastie, 1993; Hastie, Penrod, and Pennington, 1983; Winter and Greene, 2007; Ellison and Munro, 2010a). Consistent with this previous theoretical and empirical literature, findings from the mock jury study (Chapter 4), demonstrated that jurors undergo those cognitive processes throughout deliberation. The development of narratives in the conducted experiments took various forms, ranging from plausible to very bizarre reconstructions, and occurred quite frequently across all ten groups in the two experimental conditions.

While the recorded frequency of narrative construction confirmed the basic propositions of the story model, Study 2 provided further novel insights about the basis of the conducted comparison. A considerable difference was indeed observed in the frequency of narrative construction between the two conditions (thirty times in the British condition and fourteen in the Italian condition). This thesis argues that this difference cannot be considered coincidental and can be attributed to the different conditions. Plausibly, the presence of a (mock) judge had an impact on the construction of narratives, as evidenced by the analysis of the video deliberations (Chapter 4) and reinforced in consideration of the accounts of the interviewed Italian judges (Chapter 3). Interviewed Italian judges explained that, as part of what they defined as an informative, directive role, they constantly remind the lay members of the relevant/evidentiary aspects on which they should base their decisions. This suggests that, if lay members start constructing a story in order to reconstruct the trial events in a narrative format, the presiding judge will promptly lead them instead towards a more legally-oriented approach. This is what the mock judge did whenever a narrative was introduced within the Italian mock juries, so that the number of narratives constructed was reduced for two reasons: the individual juror who received the warning rectified their reasoning and, consequently, the others experiencing the situation tended not to repeat it. This obviously could not happen in the British condition, where, in the absence of a judge, the creation of a story by one juror tended to prompt more stories to be constructed in reaction, reflecting different jurors' background knowledge, bias, expectations, etc. (Pennington and Hastie, 1993; Winter and Greene, 2007). Therefore, while it appears impossible to completely eradicate this tendency of jurors towards making sense of evidence and trial facts through story construction, it can be argued that a judge's contribution can reduce the occurrence and incidence of the phenomenon.

The potential impact of the requirement for a motivated verdict was also considered as one of the explanations for the difference in frequencies of narrative construction in the two conditions. Theoretical propositions regarding the impact of motivating reasons on choices and decision-making have emphasised how the need to externally express justifications might impact the decision-making process, such that the

decision-maker will tend towards a decision that is justifiable within the group and by “superiors” (Simonson, 1989). Such a decision, in the context of jury deliberation, is one that is rigorously anchored to the entire evidentiary framework. Accordingly, it cannot be excluded that the fact that Italian mock jurors knew from the outset of deliberation that they had to publicly provide reasons for their verdicts may well have affected their approach to reasoning about the case facts throughout the entire discussion, reducing (or repressing) at times their reliance on story construction. However, on the basis of the findings (given that a certain number of narratives were still created by mock jurors in the Italian condition), the influence of the variable cannot be conclusively confirmed. Moreover, even if it operated on those jurors who did not create narratives and interpreted the evidence “appropriately”, these processes would have happened internally and could not be confirmed to have played a role.

5.3 Impact of composition and motivated verdicts on jurors’ errors

Errors related to the fallibility of human memory, the interpretation of the evidence, the understanding and application of the law have been recorded and investigated in previous research (e.g., Loftus, 1996a; Pennington and Hastie, 1988; Myers, Reinstein, and Griller, 1999). This thesis confirms previous research findings in identifying jurors’ errors in the context of the two conducted studies. In Study 1 jurors errors could be identified through the real life description of deliberations offered by Italian judges who work with jurors. In Study 2, errors of the three typologies indicated above clearly emerged in the analysis of the mock jury deliberations. This thesis offered additional insight which, going beyond the recording of the occurrence of the phenomenon, focussed on evaluating the potential impact of the two manipulated variables.

In Study 2, consistent with previous theoretical and empirical literature, memory errors were generally made because jurors’ memories were faulty at the acquisition, retention or retrieval stage, or because the uncertainty of the situation led jurors to make incorrect inferences (Loftus, 1996a; Hannigan and Tippens Reinitz, 2001). Evidence-related errors were made when jurors misunderstood and/or overestimated

evidence, which – as argued by previous literature on the topic – happened because laypeople lack the competence needed to fully understand the probative value of such complex pieces of evidence (Shuman and Champagne, 1997) and, in several occurrences, this lack of knowledge tended to be replaced by myths and misconceptions which operated in place of more rigorous parameters of evaluation. Law-related errors were caused by jurors' lack of knowledge and/or accurate understanding of the legal principles to be observed in order to reach their decision (Frank, 1973; Elwork, Sales and Alfini, 1977; Vidmar, 1994; Fisher, 2000).

The occurrence of all these types of errors was firstly acknowledged by the judges in Study 1. Throughout the interviews, the judges referred to these errors mostly in explaining how they interacted with the lay members and, therefore, how they reacted to the committed errors. When similar errors occurred in the mock juries, and the mock judge in the Italian condition reacted, as per instructions received on the basis of the interviews, the hypothesised influence of the “presence of judge” variable was confirmed. Indeed differences in the frequency of errors were recorded between the two conditions, which could be explained by the fact that the judge's warnings towards individual jurors functioned as a deterrent for the commission of new errors. Additionally, the judge's reaction to the errors was particularly impactful, as it amended jurors' memory, perception, and understanding of the mistakenly conceived piece of information.

Therefore, when a juror committed a memory error, the mock judge intervened, using the written summary report at her disposal, to dispel any doubts. Similarly, when evidence was misinterpreted or overestimated, the judge intervened, using her expertise, to suggest a correctly weighted interpretation of the disputed evidence. Likewise, when legal principles were misunderstood (as in the case of the reasonable doubt threshold, almost always conceived by jurors as 100% certainty), the judge intervened, clarifying the legal concepts. The effects of these interventions on the nature and content of the discussions cannot be ignored, as in purely lay juries, the errors generated confusion throughout the discussion, while in mixed juries, the resolution of these issues increased clarity, and benefitted the discussion.

Whether or not the need to provide a motivation impacted on jurors' commission of errors cannot be asserted with certainty. Given the nature of some errors, a connection between them and motivating reasoning can be plausibly hypothesised, in that some of the errors came as a result of the jurors' need to justify their opinions. Inferential errors were often made in order to confirm pre-existing incorrect conclusions, evidence was overestimated in an attempt to provide further support for motivating arguments, and the misuse of legal concepts was often a way to motivate and persuade fellow jurors to change their mind. All this may further demonstrate that asking jurors to think about their reasons means offering them fewer opportunities for committing such errors, when requiring them to think about their legally-rigorous reasons before committing those errors. However, these conclusions are a tentative interpretation of the findings in the light of existing literature, in that they concern internal processes that are difficult to detect with certainty.

5.4 Impact of composition and motivated verdicts on jury deliberation dynamics

The last area on which the impact of the two variables will be discussed embraces all those phenomena that occur during deliberation and that are intrinsically related to group dynamics and its effect on individual decision-makers. Social psychological research has widely investigated dynamics of leadership and conformity within groups, deepening the understanding of how groups are natural contexts wherein some individuals tend to overpower others (Pigott and Foley, 1995; Zambuto, 2016), who respond to pressures by privately or publicly conforming (Peoples et al., 2012; Smith, Mackie and Claypool, 2014). Both Study 1 and Study 2 offered original insights into the matter. Study 1 provided a perspective on within-jury leadership dynamics as perceived by those who indeed are the formally recognised leaders (Zambuto, 2016) of those groups. This, in turn, offered an opportunity for comparison of the results gained with similar dynamics at play in the mock juries used for Study 2.

In Study 1, the (real) judges never defined themselves expressly as "leaders". They emphasised the directive/informative nature of their role, explaining that they

considered themselves “the same” as the lay members, as confirmed by the fact that their votes carry the same weight. However, these views were often contradicted by the same judges, eventually showing a tendency towards influencing the lay members, by leveraging their greater knowledge and expertise. Study 2 provided further insight when the described dynamics were observed in action. The video deliberations showed that indeed the mock judge, aware of her different position, tended to embody a leading role. This resulted not only in the mock judge directing the discussion in a way that was influencing the jurors, but also in a leading tendency that helped control conflicts among jurors. This behaviour, mentioned by the interviewed judges in Study 1 and replicated by the mock judge in Study 2, generated very interesting results in the comparison of leadership perception in the two experimental conditions. The majority of the mock jurors in the Italian condition – who had a formally recognised leader – declared not to have perceived the presence of a leader, as opposed to the majority of the British mock jurors who declared to have perceived the presence of a leader. In this sense, the presence of a mock judge had an impact.

Dynamics of conformity were observed in Study 2, because they mainly regarded the lay component of the juries and, therefore, were not specifically addressed in the interviews conducted for Study 1. Data from Study 2 confirmed that conformity occurred in the mock juries, in that all groups reached a unanimous verdict, but none started from a unanimous position. More specifically, 61% of mock jurors across both conditions started with a pre-deliberation verdict preference of “not guilty” for rape, 73% started with a pre-deliberation verdict preference of “not guilty” for murder. The lower percentages (39% for rape and 27% for murder) of jurors opting for a pre-deliberation guilty verdict can be attributed to the ambiguous nature of the experimental stimuli, which reasonably led toward acquittal. Moreover, the split of pre-deliberation verdict preferences by condition for each count of indictment showed a general greater leniency on the part of Italian mock jurors: 67% (for rape) and 88% (for murder) of Italian jurors started from a not guilty verdict preference; by contrast, 88% (for rape) and 59% (for murder) of British mock jurors started from a not guilty verdict.

Given that in a pre-deliberation phase, mock jurors in both conditions were exposed to the same stimuli and no occasion for the judge to intervene had occurred yet, definite conclusions regarding this difference cannot be drawn and the potential influence of the judge should be excluded. The only explanation that can be proposed on the basis of the variables manipulated regards the need to provide a motivated verdict. Italian mock jurors were told from the beginning that they had to provide a motivation for their verdict, whereas British jurors did not receive any indication in that sense. Therefore, given the demonstrated interaction between motivating reasons and choice made (Dietrich and List, 2013), it cannot be excluded that mock jurors who knew they would be held accountable for their choice were prompted to further reflection, even in a pre-deliberation phase, and in greater percentage opted for the most reasonable choice.

In both conditions, given the final unanimous not guilty verdict reached, mock jurors who had already chosen a not guilty verdict did not need to change their mind in order to adhere to the group decision. All the other mock jurors had to conform and the analyses conducted showed whether they privately (internally changing idea) or publicly (externally adhering with the group, but keeping their own idea) conformed. According to the findings, across the two conditions and for both counts of indictment, private conformity occurred less often than public conformity. This is consistent with previous literature, according to which people are more willing to formally comply with the most desirable opinion than to change their minds internally (Kelman, 1958; Kessler, 1975). Moreover, mock jurors in the Italian condition tended to conform privately more often than mock jurors in the British condition, to the extent that, for the murder accusation, all the Italian mock jurors who changed their opinion during deliberation, did so by privately changing their mind. In contrast, the British mock jurors who conformed did so mainly publicly.

In this respect, presence of the mock judge and requirement for motivated verdicts in the Italian condition are thought to have had an impact. As was noted also during the video deliberations, Italian mock jurors constantly received answers and clarifications to their questions and doubts, and were prompted to reflect on the evidence when

asked for their (evidence-based) motivated verdict. It is not surprising thus that the resolution of doubts throughout deliberation and the increased confidence in better understood probative value have led mock jurors who changed their verdict preference to do so because they were convinced that the group's opinion was correct, as is typical of private conformity (Smith, Mackie and Claypool, 2014). Reasonably, if convincing people that the group's opinion is the most accurate and socially (or rather, legally, in this case) acceptable is the most effective way to lead them to privately conform, there is no doubt that this is exactly what the mock judge has done throughout the deliberations. Accordingly, British jurors who adjusted their opinions to reach a unanimous verdict, not having the judge's aid, presumably brought their doubts with them and therefore were not willing to change their mind privately, although they felt compelled to do it publicly.

Nonetheless, this did not amount to an increase in jurors' overall confidence with their verdict choice. As seen, confidence levels were not significantly impacted by differences in the experimental conditions. However, the meaning of this result should be carefully considered to not undermine the rest of the findings. Jurors will, in any occurrence, know that they are determining another person's fate, and, given the difficulty, seriousness and uncertainty of the to-be-judged situation, high levels of confidence should not be expected whatsoever, especially when the evidence presented is weak and ambiguous. Asking jurors to be confident with their verdict preferences requires them to be convinced that the defendant has or has not actually committed a crime. Different is to ask jurors about the fairness of the rendered verdict. A fair decision, in the context of a criminal trial, is the legally correct verdict, which may be in conflict with the factually correct decision. For example, the legally correct verdict might be 'not guilty' (e.g. the case was not proven beyond a reasonable doubt), yet the defendant may well be factually guilty of the crime. Therefore, even a juror who is not confident in their verdict preference (i.e. not sure the defendant committed the crime) may well consider that verdict fair (i.e. correspondent to the legal standards).

Indeed, consistent with conformity tendencies, mock jurors' perceptions of verdict fairness demonstrated that Italian mock jurors found the group verdicts fairer on average than the British jurors. This finding is reasonable considering that Italian jurors privately conformed more than British jurors, showing that they internally believed in the fairness of the verdict. Similar considerations to those reported above have to be proposed in this case with regard to the impact of the manipulated variables on the recorded statistically significant difference between conditions. The prompt to reflect on motivating reasons along with the constant remarks that the mock judge proposed on the importance of legal principles (burden of proof, presumption of innocence, reasonable doubt) and evidence-based evaluation of the case, evidently increased jurors understanding of what constitutes a "fair" judicial decision.

5.5 Lay vs Mixed Juries – is there a “winner”?

The comparison reported in this thesis allowed an observation of the effects of a presence of judges on juries. To enhance the results already reported, further reflections on this specific aspect are due. The usefulness of the presence of a judge on the panel was evaluated by mock jurors in Study 2. Findings showed the existence of a significant difference between British and Italian mock jurors in their respective perception of the need for and the usefulness of a judge or someone to direct the discussion. British mock jurors showed a lesser tendency to feel the need for either a judge or someone to direct the discussion as opposed to Italian mock jurors. These findings can be explained as the result of a greater awareness on the part of Italian mock jurors, who actually experienced the presence of a judge, and therefore of someone to direct the discussion, and the benefits that such a presence generated.

However, in light of the results of Study 1 and after careful considerations of positive and negative aspects of the presence of judges on jury panels, it is clear that the situation is not straightforward. Study 1, specifically digging into the role of judges who work on juries has unearthed a number of problematic aspects that having a judge on the jury may cause, and Study 2, by showing a (mock) judge in action, confirmed those findings. First of all, a series of contradicting remarks emerged from the interviews of judges, revealing how some aspects of their role were different in practice from the

description they provided on the basis of their perception of them. The most salient aspect of contradiction was found in fact that the judges defined their role as neutral and informative, while in reality they appear to influence the lay members. Study 2, through the behaviour adopted by the mock judge (who only received general instructions on how to act, but was not expressly asked to influence the jurors) and the reaction of mock jurors to that behaviour, offered an opportunity to observe some of these dynamics in action. The mock judge always started the discussions with a quite balanced, neutral approach, never providing her opinions first and asking jurors what they thought, ensuring that each of them had a chance to talk. However, during the unfolding of deliberation, the mock judge, being influenced by the heat of the discussion was not able to maintain a balanced demeanour and ended up intervening quite frequently. While at times her interventions were aimed at clarifying legal aspects and/or misunderstandings, and were therefore beneficial, other times they appeared to result in subtle techniques to persuade the discussion in the “right” direction.

In light of these considerations, the overall beneficial nature of the presence of these professionals on mixed juries comes under scrutiny, especially when considering it in conjunction with the very nature and essence of the trial by jury. A system that was created on the premises of the need to ensure popular values into the judicial process (Di Majo, 2014) appears barely compatible with such mechanisms of influence from actors connected to the State power. Indeed, as Moschella (2009) compellingly argues, it is undoubted that the presence of lay members on the Italian jury “forces” the professionals (the judges) to a different evaluation of the matter, through a process that merges the legal/technical competence of the judges with the community spirit and population perspective of the lay members, which overall should grant greater independence to the judgement. However, it is probably too ambitious to believe that such a system would accomplish an actual direct exercise of the judicial power on the part of the citizens. In fact, as observed, ‘the judges on the jury, due to their knowledge of the procedural techniques and to their experience, have an easy game in imposing their opinion on the jurors, unavoidably influencing the jurisdictional decision’ (Moschella, 2009, p.160, *translated from Italian*).

And yet, that is only one of the problematic aspects that the mixed jury, as designed in the Italian system, presents. If it is true that the need from which the jury system stems is to reinforce the democratic legitimisation of the judicial process through the presence of citizens in criminal trials, some incongruities are difficult to understand and need resolution. It is, for example, hardly comprehensible the rationale behind the mechanism of appeal, whereby Italian verdicts rendered by mixed juries are then referred to panels of (only) professionals, if appealed (Pizzorusso, 1990). This means following the same procedure to which other judicial decisions are subjected and, in so doing, depriving the lay component of the Italian jury of any power whatsoever. This is even more worrisome, considering that, as was noted and will be further addressed in the following section, according to the Italian system, the reasons for appeal have to be grounded in the verdict motivation, which is entirely written by the judges (and not by the jurors).

The existence of such deep contradictions leads to reflection on whether there is any actual intention to involve the community in the direct exercise of justice, as well as to question the reliability and effectiveness of a mixed jury system no less than a purely lay one. It is clear that both alternatives present flaws and leave themselves open to strong criticism. Purely lay juries have long been blamed for the lack of competence and rigor that underlie decisions made by inexperienced laypeople. By contrast, mixed juries, in order to guarantee a degree of competence and rigour to those decisions, in fact appear to deprive the system of some of its most fundamental characteristics. In a context where neither of the two systems appear to satisfactorily address justice needs, solutions must be found. With this intent, suggestions of resolution and further research agendas are proposed at the end of this chapter.

5.6 Guilty or not guilty – what if we asked ‘why’?

Another difference that was at the heart of the comparison in this thesis is the requirement for Italian mixed juries to provide a motivation for their verdicts, which is not required by British lay juries. This thesis used the term motivation in two different senses: “motivation” is the written document that mixed juries have to provide as part of any verdict (and any judicial decision) in Italy, and motivation is also the actual

reason that jurors, abiding by the formal requirement, provide for their choices. Study 1 provided information from the perspective of those judges that have to write the motivation (legal document), and Study 2 added to that information, providing further insight into the actual reasons behind verdict decisions.

Considering the required motivation, Study 2's findings revealed interesting differences between the mock juries who were not expressly required to provide a motivation (British) and those who had to do so (Italian). When analysing the nature of the motivations provided, all Italian mock jurors offered legally-oriented motivations, while only 63% of British mock jurors did so. This statistically significant difference can be attributed to the different deliberation conditions, whereby those mock jurors who knew they had to legally motivate the verdict started thinking about that from the beginning. The benefit of such a requirement, in this sense, is clear: it modifies the entire approach of jurors towards deliberation, presumably informing the entire reasoning process and correctly orienting it toward the legal aspects of the case. This does not prevent mock jurors from engaging in expected human cognitive processes (narrative construction, confirmation bias, etc.) or to make common errors, however it might limit the damage of such approaches, by leading the mock jurors on the right track every time they are required to think about their reasons.

Regarding the motivation conceived as a legal written document, the information obtained from the judges interviews partially reiterated and confirmed the rules set by the Italian criminal justice system (what it is, who has to write it, etc.). It also revealed judges' dislike towards such a long and detailed document, the writing of which is perceived as a burden. Interestingly, judges explained how they write a motivation with which they do not agree. Since, as judges unanimously recognised, given the strictly legal nature of the verdict motivation document, it is their job to write it because mock jurors could not, the possibility of them having to motivate a verdict with which they do not agree could not be discarded. The instance was considered possible, in that lay members are the majority in Italian juries and, at least theoretically, should be able to reach a verdict that is contrasting with the judges' opinion. However, if choices are connected with motivating reasons, so that

arguments and conclusions are eventually mutually reinforcing (Mercier and Sperber, 2011), it is difficult to imagine how one can convincingly argue in favour of something in which they do not believe. In this context, the phenomenon of the “suicidal sentences” was mentioned and described as a tactic that judges may use to react to the need to motivate a verdict they would not opt for. By writing a motivation that is intrinsically “vitiating” (i.e. flawed, incomplete, incoherent), the judges obtain the result that it will be appealed.

An appealed motivation goes into the standard judicial process that does not include laypeople, depriving the public of any involvement in the process. The motivation system, in this sense, rather than being a democratic guarantee – as it should be in line with the purpose of the jury trial and of the overall Italian judicial system – results in a further instrument of control of the professional component in jury trials. Therefore, if prompting jurors to think about their reasons and to ensure they are legally acceptable has its benefits, on the other hand, this instrument – as it is currently used – ends up being detrimental to an exercise of the popular sovereignty and requires considerable amendments in order to effectively work.

5.7 Contributions, strengths, limitations and future research suggestions

The existence of extensive theoretical and empirical literature on the topic of jury decision-making is easily explainable in light of two reasons. Firstly, the jury task represents a significantly illustrative example of how people, individually and collectively, make choices. Secondly, the impact of this unique task exceeds a pure scholarly dimension, given the seriousness of its implications in legal matters (Hastie, 1994). In an effort to better understand the issues at the heart of the (mal)functioning of the jury trial, extensive research has focused on the cognitive processes and dynamics that underlie decision-making. This thesis makes an original contribution to the previous literature on the topic by shedding light on the impact that procedural aspects of the jury trial may have on those cognitive processes, by comparing two different systems’ functioning. This comparison helped to confirm previous findings

and to add novel insights about the intricacies of the jury trial. In doing so, the strengths of this thesis were accompanied by limitations, which in some cases could not be avoided and were therefore acknowledged, particularly with consideration of their potential effects on the findings. Accordingly, an account of strengths and limitations is provided in this section.

This thesis, starting from the assumption that no criminal justice system is “perfect”, made use of the benefits of comparison. Indeed, ‘an examination of other systems... can help to illuminate each system, cause us to reflect on how it deals with problems, and hopefully generate ideas about how it can be made better’ (Vidmar, 2000, p.52). Therefore, a comparative study never benefits only one country or another (as if there were a “worse” and a “better” system), yet through insights into another reality, it creates an exchange from which both countries might benefit. With this aim in mind, the researcher’s connection with both the Italian and the British criminal justice system was employed in designing and conducting a research project that might achieve the ultimate goal of suggesting solutions to a problem that, in different ways, affects jury trials in both countries.

Fundamental in this respect was the opportunity to approach and interview real Italian judges, who provided very valuable insights and unique perspectives on the actual functioning of the jury trial behind the closed doors of the deliberation room. In a context where British jury deliberations cannot be accessed and British jurors cannot be interrogated (Ellison and Munro, 2013), as well as Italian deliberation rooms cannot be entered, obtaining Italian judges views constituted a considerable and novel advantage of this thesis, and one of the key aspects of the conducted comparison.

The involvement of Italian judges in this research contributed in several ways to the original contribution that this thesis makes, in that not only findings from the interviews generated further understanding of the internal functioning (and intrinsic flaws) of the Italian jury trial, but they also stimulated further reflections on benefits and limitations of the presence of professionals on jury panels as well as on how their role could be shaped in view of an overall improvement of the jury system.

Accordingly, an implementation of the findings into practice would suggest the creation of a new figure, a “professional juror”, who should undertake appropriate training, to adequately exert their role. On the basis of what emerged from the findings of this research, the training should embrace several areas in order to constitute a counteractive measure to avoid most of the problematic aspects emerged from the investigation of the judges behaviour.

Firstly, the “professional juror”’s role should be precisely defined to prevent it from being overly discretionary, and to ensure that it will be exerted within clearly set boundaries. The “professional juror” should be, therefore, made aware of the purpose of their role of guidance, warned to refrain from exerting any influence on the lay jurors, instructed on specific issues of relevance in criminal trials (e.g. interpretation of forensic evidence and/or probabilistic evidence, of which the interviewed judges admitted to lack understanding). Such training would contribute to create a professional who, by possessing the necessary competence in legal matters and by knowing how to properly exert their directive role, would resolve the main issues identified so far with the jury trial: respectively, lack of competence on the part of the lay jurors, and overly discretionary power into the hands of the judges. Future research should be conducted on this aspect, with the aim to test the effectiveness of such training and its ability to achieve its main goals.

Continuing with the sample used in this research, the use of mock juries has to be considered, for different reasons, as both a strength and a limitation. General limitations associated with the use of mock jury simulations have been widely addressed by previous research and highlighted in this thesis. They mainly focus on the unrealistic nature of such experiments and on the effects that it may have on the findings given that participants cannot experience the seriousness of real jury duty responsibilities. These observations are all sensible and were taken into account in this thesis. However, the choice to use this experimental tool, despite its pitfalls, was based on some generic and specific considerations. Broadly speaking, similar to previous research endeavours, the mock jury paradigm was considered the best solution given the restrictions forbidding real jury research. Specifically to this thesis, the fact that the

design of the mock jury study was strategically interconnected to the judges' interview study enabled it to benefit from the strengths of triangulation, through which its findings could be confirmed and reinforced. Moreover, the expected and feared effects of the lack of seriousness allegedly ascribable to the simulated nature of the task were not observed in the current study. As in other studies (Chesterman, Chan and Hampton, 2001; Ellison and Munro, 2010a), mock jurors appeared very engaged, seriously focused on their decision, and absolutely mindful of the gravity of its implications, as was demonstrated by the numerous claims that jurors made regarding feeling stressed and finding the whole situation quite difficult to handle. All things considered, it was concluded that the benefits of observing actual interactions between human participants (whose recruitment indeed constituted another challenge) in a study investigating human cognitive processes would overpower the methodological issues that naturally come with any research simulations.

Participants' recruitment was a further issue in both studies. In Study 1, only three judges agreed to be interviewed. As explained in Chapter 3, this was not an overly problematic aspect, given the purely qualitative nature of that study and considering the advantages that the judges' perspective offered. However, a greater number of judges would have further increased the richness of the data. In particular, the judges were purposely selected from different courts, and recruitment on a larger-scale would have improved the study, providing further perspectives which could have further confirmed or disproved those acquired. Especially in consideration of the importance of the judges' accounts on the matter, further research should achieve the goal of conducting the investigation on a larger geographical scale.

A number of problematic aspects concern the sample recruited for Study 2. The majority of mock jurors in the British condition were students. The use of students in mock jury studies, despite being very common, is criticised due to the unrepresentativeness of such participants (Hastie, Penrod and Pennington, 1983; Diamond, 1997). However, research has shown that it may be the case that demographics of the sample does not significantly affect jury deliberations (Bornstein, 1999). More recent research has further confirmed this finding, by demonstrating that

some specific aspects of deliberation (guilty verdicts, culpability ratings, etc.) do not vary with the sample (Bornstein et al., 2017). In addition, some argue that, even though students may be not an ideal sample, if they are the only alternative to other even less reliable methods, then their employment in mock jury research is considered acceptable (Nuñez, McCrea, and Culhane, 2011). Indeed, an alternative to students could have been online research, which would have produced higher participant numbers and include a wider variety of participants from the general public. However, excluding deliberations to have individuals answering pre-set questions on their own would have affected realism in much higher measure. Having people, even though they might be students, engaging in deliberations that allow to observe the spontaneity of a free-flowing discussion has to be considered preferable to the absence of any jury deliberations (Ellison and Munro, 2010a; Nuñez, McCrea, and Culhane, 2011).

A more problematic aspect with the students used for this research is that they were predominantly Criminology students. This resulted from opportunity sampling, and was considered as potentially influential, in that those students were expected to have a background knowledge of legal matters that would make them quite different from random members of the public. This limitation certainly needs to be addressed by future research using a different sampling strategy. However, without neglecting the impact that this element may have had on the results, for the purpose of the analyses presented in this thesis, it is worth noting that the presumed higher-degree knowledge of Criminology students did not prevent them from discussing the issues and/or making errors that are found to be made by members of the general public too. Therefore, their contribution was found to be useful and was, with due caution, analysed.

The opposite was true in the Italian condition, where, given the lack of connection of the researcher with the university environment in Italy, students were not the most approachable sample, and therefore the majority of mock jurors were members of the general public. This circumstance, despite being closer to the requirements of representativeness that research standards strive for, was considered problematic due

to the resulting mismatch that it generated with the other experimental condition. This issue was not disregarded either, yet its potential effects were considered in light of previous research that suggested participant demographics are not found to affect mock jury research (Bornstein, 1999). Despite the existence of research that argues against these methodologies, claiming that more research on the matter is needed (Wiener, Krauss and Lieberman, 2011), findings from this study did not detect any differences in this sense. Italian mock jurors addressed the same matters, made similar mistakes and reached the same verdict, to the extent that a close comparison was possible to make. Notwithstanding this, greater uniformity and representativeness of the sample is certainly advisable for future research endeavours.

A related problematic aspect that regarded the demographic element in the experiments was participants' age. As specified in section 2.4.2, there is a different minimum age requirement in real-life British and Italian juries: eighteen in the UK, thirty in Italy. The higher age required for Italian juries seem to somehow reflect a tendency of the Italian criminal justice system to attribute to higher age greater decision-making abilities. This is observable also in the rule whereby, in Italian juries, to avoid influence, votes start from the youngest juror (CPP, art. 527, para. 2). However, this tendency is not clearly explained, and appears in fact grounded on the common rule of thumb whereby aging brings with it experience and wisdom. Starting from this assumption and given the inconclusive results that previous research reached in terms of whether demographics affect deliberations, this particular aspect of difference between Italian and British jurors was not specifically analysed in the experiments that this thesis presents. However, this could be a further interesting object of analysis in future research which could specifically address the aspect of demographics and investigate whether or to what extent they might affect deliberations.

The experimental stimuli for the mock jury experiments was designed using an existing fictional case developed as a university teaching tool and amending it according to the purpose of the study. The fictional case presented was intended to be ambiguous and controversial, in order to prompt extended deliberations. As demonstrated by the rich

amount of data emerging from the study, the goal of stimulating discussions was satisfactorily achieved. However, the way the case was designed and the weak nature of the evidence included appear to be the cause of the uniformity in verdicts (all mock juries acquitted), which did not allow any conclusions to be drawn about the potential impact of the variables under analysis on final decisions. This issue has occurred in other studies (e.g., Woodhams and Labuschagne, 2012), and was foreseen and accepted, in consideration of the fact that the opposite problem would have been more detrimental. If the case had contained pieces of evidence strong enough to lead to a conviction, the risk was to have juries that focused on that only and did not address other matters, which would have jeopardised the study.

However, one of the measures taken to ensure that the foreseen uniformity of verdicts did not result in hiding or suppressing important individual's perspectives, was the use of pre- and post-deliberation questionnaires, aimed at measuring individual jurors' views and verdict preference. The benefits of this methodological tool, used in conjunction with the video deliberations, have been emphasised throughout this thesis with reference to their specific functions in the context of this research. These benefits were evident in that they offered a multi-faceted view of the phenomena at play, so that each of the elements emerging from the questionnaires could be verified by the videos to see whether there was correspondence between mock jurors claims and perception about the same phenomenon. The pre-/post-deliberation questionnaires represented, therefore, a key instrument to obtain an account of jurors' individual insights and to overcome the by-product of the uniformity in verdicts outcome.

Findings from the studies described in this thesis contribute to the existing body of literature on jury decision-making, by adding original insights into the impact that the variation of procedural aspects may have on jury deliberations. Of the two variables observed, the presence of a judge had an impact on deliberations, affecting jurors' behaviour and understanding at different stages of the discussion; the requirement for jurors to motivate their verdicts showed its impact on the nature of the motivating choices as reported at the end of deliberations. Both studies and related findings have, therefore, contributed to increased understanding of jury decision-making and the

impact of procedural variation on such decision-making. However, a further all-comprehensive consideration of the findings might also stimulate reflections on their potential applicability beyond the realm of the jury trial. Indeed, observing the findings from a different angle, their application to judicial decision-making in magistrates courts appears sensible and prospectively beneficial.

Most information, acquired through the Italian judges interviews and the mock jury experiments, on judges' discretionary decisional power, decision-making mental processes, lack of understanding and critical awareness of forensic and probabilistic evidence, is presumably applicable by extension to magistrates in English/Welsh courts likewise. Without neglecting the need for further investigation into the matter before drawing definite conclusions, it seems plausible to believe that judges undergo similar decision-making processes, facing similar challenges in the exercise of their role. It is in this sense that findings from this research may represent points of reflection for further development and potential improvement of the system as a whole. For example, requiring the magistrates to provide motivated judicial decisions might result into greater guarantee and uniformity of decisions, which in turn might increase the fairness of trials. Likewise, providing magistrates with greater knowledge and understanding of forensic and probabilistic evidence might be a way to also confer them greater critical awareness towards expert witness testimony for it not to be blindly believed. These suggestions are proposed here as a stepping stone from which further research should be designed and undertaken in order to tailor future investigations on the specific needs that might arise in this context.

To conclude on the specific applicability of this research findings to the context of the jury trial, it appears clear that the conducted comparison was useful in identifying solutions for the problematic aspects that the jury trial currently presents and to suggest what an "ideal" jury system might look like. British and Italian systems represent the two extremes, from which beneficial aspects of both systems may be combined to obtain a more rigorous jury trial. If, on the one hand, the lay component of the jury trial should be maintained as an indispensable characteristic of the very essence of this institution, on the other hand, purely lay juries, left uncontrolled and

without direction, may continue to create issues. By contrast, the use of judges with overly discretionary power and the possibility to largely influence lay members and/or deprive their opinions of any strength equates to overwhelming the lay component, which should be not acceptable. However, saving the most desirable features from these two systems and amending them may lead to an ideal scenario in which lay people deliberate, uninfluenced, under the guidance of a figure (a “professional juror”), that is purposely trained to only do what is actually needed to put jurors in the best position to make a sensible, well-informed, legally-oriented decision.

Overall, this PhD research has contributed original research to existing knowledge on the topic of jury decision-making. However, a number of limitations should be addressed in future research to further consolidate the validity of its results. Therefore, in light of the undeniable relevance of the topic in the context of the criminal justice system and on the basis of the mainly exploratory nature of the findings of this thesis, future research is crucial to continue increasing the understanding of decision-making in the context of the jury trial and propose further solutions.

CONCLUSION

Jurors – in the words of the jurist Glanville Williams, ‘a group of twelve men of average ignorance’ (cited in Lloyd-Bostock, 1988) – are considered not adequately equipped to properly assess complex evidence and understand the law (Kassin and Wrightsman, 1988). However, democratic principles permeating criminal justice systems require the injection of community values in the exercise of justice. Therefore the lay perspective that ‘averagely ignorant’ jurors confer to trials appears, nonetheless, needed. On the other hand, greater awareness of legal principles and rules, such as that possessed by professionals working in the legal field, would allow for legally fairer decisions. The tension that these opposite needs create is at the heart of the debate around the jury trial.

The studies presented within this thesis have provided an original contribution to the debate, offering further insights into the jury trial, through the comparison of its functioning in two countries, England and Italy. Comparisons create beneficial exchanges that deepen the knowledge of both the compared elements (Vidmar, 2000), therefore this thesis has considered differences existing in the two jurisdictions. Jury composition (presence/absence of professionals on the panel) and requirements for verdict type (motivated/unmotivated verdicts) have been the variables manipulated and observed in this thesis to understand the impact of these factors on jury decision-making.

Results obtained from interviews with Italian judges and mock jury experiments suggest that both the above-mentioned factors have effects on jury decision-making. Impact of jury composition and motivated verdicts was found on juries’ behaviour, errors, and deliberation dynamics. Specifically, it was found that the presence of a (mock) judge on jury panels offered a number of benefits, by ensuring consistency among juries’ choices (e.g. in the choice between verdict-driven or evidence-driven approach to deliberations); reducing the frequency of narrative construction; preventing and/or correcting jurors’ memory-related, evidence-related, and law-related errors; exerting leadership, hence controlling emerging conflicts among jurors; and inducing more frequent private than public conformity along with a higher degree

of perceived fairness of verdicts on the part of the (mock) jurors. In addition, the requirement for motivated verdicts was beneficial in reducing the occurrence of confirmation biases in jurors' reasoning, as well as contributing to benefits produced by the other variable in several aspects, e.g. reducing narrative construction.

Fundamentally, the presence of a (mock) judge has been demonstrated to provide the jury trial with those elements of professionalism and legal competence that are often identified as lacking (and, arguably, needed), so that jurors' 'ignorance' was filled by the intervention of a judge. On the other hand, the requirement for motivated verdicts has been seen as adding further guarantees, by imposing greater rigour to jurors' reasoning, so that even if they were naturally prone to make a non-legally-oriented decision, the awareness of the need to provide an acceptable justification prompted further reflection on their part.

However, findings from the studies presented in this thesis also demonstrated that the presence of a judge on jury panels frequently exceeds its directive and supportive nature and can be overly influential, which becomes problematic in view of the original intentions on which the jury trial has been founded. Professionals' influence, exerted to the extent of depriving jurors of their freedom and decisional power, contradicts the very nature of a trial that intends to make use of community values to grant democracy. Consequently, a jury trial, only formally left into the hands of the citizens but substantially retained by the State power, appears as "dangerous" as one where citizens are left alone.

In light of these findings, it is clear that neither of the two systems currently proposes the most effective solution for the functioning of the trial by jury. Being at the opposite ends of the spectrum, the British and the Italian systems encounter different, but equally worrisome, issues. All things considered on the basis of this comparison, an effective way to attempt solving these issues might be to create an ideal scenario which integrates benefits and overcomes limitations of both systems: a system in which jurors can deliberate uninfluenced, but being guided by someone who, trained to only guide and help them, would ensure that the decision-making process leads to a well-informed, legally-oriented verdict.

Most previous research on the topic of jury decision-making has attempted to identify problems and, consequently, find solutions for the limitations of the jury trial presents. The great deal of discretionary power that juries are assigned in determining another person's fate, along with the disturbing figures of incorrect jury verdicts, make this an important focus of robust research. Given the fact that juries are likely to continue to feature in criminal trials for the foreseeable future, research on the topic should continue to concentrate its efforts on proposing potential solutions to improve the adequacy and legal fairness of jury decisions. This thesis makes an original contribution to the literature on the topic, and it is hoped that further research will continue to produce an increased understanding and improvement of jury decision-making, in order to move towards the ultimate aim of reducing miscarriages of justice and grant a truly legally fair trial by jury.

APPENDIX A

Italian judges' interviews – Informed consent form

(Translated from Italian)

A STUDY ON JURY DECISION MAKING

INFORMATION SHEET

Introduction and Purpose of the Research

This research project is conducted by Cristina d'Aniello, a Ph.D. student at the University of Leicester – Department of Criminology, and supervised by Dr Lisa Smith.

You are being asked to participate in a research investigating the topic of jury decision-making. The aim of this study is to assess how juries make their decisions when reaching a verdict in criminal trials. Furthermore and more specifically, this study will be intended to better understand the role of professional judges who sit on juries (in some countries, such as Italy) and whether or to what extent their presence on the jury panel may play a role in the entire decision-making process. It is believed that to obtain a fuller understanding of the various factors that affect jurors' decision-making might be a rewarding result to reach, in order to help avoid the occurrence of miscarriages of justice ascribable to wrong jury verdicts. Your participation may, therefore, help clarify to what extent it is possible to achieve this result and, although you may not directly benefit from this research, findings from the study may be useful for your community and for society as a whole.

Participation

You have been selected as a participant in this study because of your professional qualification and job position, which make you particularly suitable for participating in this research. Indeed, given the purpose of the study, only judges who work (or have worked) on jury panels may be able to provide accurate information about the characteristics of their role in those contexts.

Your participation in this research is entirely voluntary. It is your choice whether to participate or not. You may also change your mind later and stop participating, even if you agreed earlier, so that if at any time during this study you wish to withdraw your participation, you are free to do so without prejudice.

Your participation in this study will not cause any risk of harm to you, since the required activity (see below) does not present the potential to generate any particular distress or discomfort and will not involve sensitive or personal aspects. You will not incur any costs as a result of your participation in this study.

Activity

If you agree to participate, you will be requested to answer some questions about the role you have when sitting on a jury in criminal trials. None of the questions will refer to any real cases that you have tried in the past and you will not be required to refer to people or specific real facts. The questions will be rather general and will only refer to the tasks you usually undertake in the execution of your duty. The interview will be audio-recorded, with the only purpose of granting a greater accuracy in the data collection. However, your consent will be asked before the interview begins. The completion of this activity should take approximately forty-five minutes.

Privacy

With the aim of respecting your privacy, a strict confidentiality will be guaranteed. You will not be putting your name on anything except this form. The information collected from this research project will be kept private and no one apart from the researcher will have access to them.

Debriefing (Sharing the Results)

The data collected for this study will be used for a Ph.D. research project. Once the data is analysed, a summary of findings may be shared more broadly, so that other interested people may learn from the research. Only the results will be reported and it will not be possible to identify any individuals who participated in the study. Furthermore, a summary of findings will be also available for you from the researcher on request.

Who to contact

If you have any questions before the interview starts, you can ask them to the researcher in person. Also, if you have any questions at any time during the study or later, as well as if you wish to withdraw your participation in this study, please do not hesitate to contact the researcher:

Cristina d’Aniello

e-mail: cd228@leicester.ac.uk

AUTHORIZATION

1. I have read the foregoing information.
2. I have had the opportunity to ask questions about the study and any questions I have asked have been answered to my satisfaction.
3. I consent voluntarily to be a participant in this study.
4. I am aware that I may withdraw at any time, without giving any reason and without prejudice.
5. I am aware that I may contact the researcher and ask questions about the study at any time.

Please retain this copy for your records

APPENDIX B

Italian judges interview questions

(Translated from Italian)

- 1) How long have you been working as a professional judge on jury panels?
- 2) Since yours is a particular role in the Italian criminal justice system (in all the other cases, even when judges work on a panel, they work in collaboration with other professionals, not laypeople), have you received any particular training/instructions on how to carry out this particular duty?
 - a) If 'yes':
 - Could you tell me more about these training/instructions? (What have you learnt from them?)
 - Were you told what the reason for your presence on the jury panel is?
 - Did you receive any instructions about how to behave towards the jurors (laypeople) throughout the deliberation process?
 - Did you receive any specific instructions about how to avoid influencing the jurors?
 - Were you told whether or not you are allowed to remind the jurors of relevant facts of the case, if they seem to have forgotten them during the discussion?
 - Were you told whether or not you are allowed to better explain facts or evidence of the trial (e.g. forensic evidence), if the jurors seem to be misinterpreting them during the discussion?
 - b) If 'no':
 - What do you think is the reason why a jury requires the presence of professional judges?
 - Given the lack of training/instruction, do you follow any self-imposed 'rules' when you interact with the jurors (laypeople) throughout the deliberation process?

- Do you usually try not to influence the jurors or, on the contrary, do you think they need to be somehow directed/controlled as they are not law experts?
- If during the discussion jurors seem to have forgotten some relevant facts of the case, do you remind them of those?
- If during the discussion jurors seem to be misinterpreting facts or evidence (e.g. forensic evidence), do you try to better explain them?

3) Do you have a direct, active participation in the discussion about the case during the deliberation? In other words, do you express your opinions at any time as well as the other jurors (laypeople)?

4) How do you perceive your role overall? Do you perceive yourself as very active in the discussions, or as a more passive participant?

5) According to the Italian criminal justice system, juries have to provide a motivation for their verdicts. Are jurors asked about the reasons of their choices multiple times – every time that it appears necessary – during the discussion? (For instance, if a juror believes that a single piece of forensic evidence strongly proves that the defendant is guilty, is he/she promptly asked why?). Or are they only asked about their motivation towards the end of the entire deliberation process?

6) At the end of the deliberation, after each member of the jury has voted and the verdict has been reached, you have to write the judicial sentence containing both the verdict and the motivation. The motivation has to show the existence of a logical nexus between the interpretation of the evidence presented at trial and the final decision. How do you keep track of all the information you need to include in this report in order to show the existence of that nexus?

7) Do you believe that lay jurors could deliberate without the presence of professional judges on the panel?

8) Do you believe that lay jurors could deliberate without providing a motivation?

APPENDIX C

Mock jurors – Informed consent form

INFORMATION SHEET

Introduction and Purpose of the Research

This research project is conducted by Cristina d’Aniello, a Ph.D. student at the University of Leicester – Department of Criminology, and supervised by Dr Lisa Smith.

You are being asked to participate in a research investigating the topic of jury decision-making. The aim of this study is to assess how juries make their decisions when reaching a verdict in criminal trials. It is believed that to obtain a fuller understanding of the various factors that affect jurors’ decision-making might be a rewarding result to reach, in order to help avoid the occurrence of miscarriages of justice ascribable to wrong jury verdicts. Your participation may, therefore, help clarify to what extent it is possible to achieve this result and, although you may not directly benefit from this research, findings from the study may be useful for your community and for society as a whole.

Participation

Your participation in this research is entirely voluntary. It is your choice whether to participate or not. You may also change your mind later and stop participating, even if you agreed earlier, so that if at any time during this study you wish to withdraw your participation (for instance, if you feel uncomfortable with the topic of the discussion), you are free to do so without prejudice.

Your participation in this study will not cause any risk of harm to you, since the required activity (see below) does not present the potential to generate any particular distress or discomfort and will not involve sensitive or personal aspects. You will not incur any costs as a result of your participation in this study.

Activity

If you agree to participate, you will be requested to take part in a mock jury deliberation; you will be part of a small group and act as a juror/judge on a fictional crime case scenario which will be presented to you (please be aware that the crime case can contain pictures and information of a violent and/or sexual nature, however, please also consider that everything you will see is fictional and no real case's images will be used for the experiment). Like in a real jury, you will need to discuss the case with the other members of the jury and reach a verdict. Before and after the group deliberation, you will be given a short questionnaire, aimed at gaining a better understanding of various aspects of your deliberation experience. The deliberation will be video/audio-recorded, with the only purpose of granting a greater accuracy in the data collection. However, this material will be only used for research purposes and will be stored securely. The completion of these activities should take approximately one hour.

Privacy

With the aim of respecting your privacy, a strict confidentiality will be guaranteed. You will not be putting your name on anything except this form. The information and material collected from this research project will be kept private and no one apart from the researcher will have access to them; no one else will hear the recordings and/or watch the videos, which will be destroyed after completion of the research.

Debriefing (Sharing the Results)

The data collected for this study will be used for a Ph.D. research project. Once the data is analysed, a summary of findings may be shared more broadly, so that other interested people may learn from the research. Only the results will be reported and it will not be possible to identify any individuals who participated in the study. Furthermore, a summary of findings will be also available for you from the researcher on request.

Who to contact

If you have any questions before the experiment starts, you can ask them to the researcher in person. Also, if you have any questions at any time during the study or later, as well as if you wish to withdraw your participation in this study, please do not hesitate to contact the researcher:

Cristina d'Aniello

e-mail:

AUTHORIZATION

1. I have read the foregoing information.
2. I have had the opportunity to ask questions about the study and any questions I have asked have been answered to my satisfaction.
3. I consent voluntarily to be a participant in this study.
4. I am aware that I may withdraw at any time, without giving any reason and without prejudice.
5. I am aware that I may contact the researcher and ask questions about the study at any time.

Name of the Participant: _____

Signature: _____ Date: _____

Name of the Researcher: CRISTINA d'ANIELLO

Signature: _____ Date: _____

Please retain this copy for your records

APPENDIX D

Mock jury experiments flyer

GUILTY OR NOT GUILTY? YOU BE THE JUROR!



PARTICIPANTS NEEDED FOR RESEARCH IN JURY DECISION-MAKING

- As a participant in this study, you will be asked to sit on a mock jury; you will be part of a small group and act as a juror on a fictional crime case scenario (i.e. a murder trial). Like in a real jury, you will need to discuss the case with the other members of the jury and reach a verdict.
- Before and after the group deliberation, you will be asked to complete a short questionnaire on your deliberation experience.
- The completion of these activities should take ca. one hour.

In appreciation for your time, you will
receive: a 5£ Amazon Voucher 

For more information about this study and/or to volunteer
please contact:

Cristina d'Aniello
cd228@le.ac.uk

PhD Student/Graduate Teaching Assistant
Department of Criminology
University of Leicester



UNIVERSITY OF
LEICESTER

Criminology - The Priory - 124 Upper New Walk
Leicester LE1 7QJ
T: +44 (0)116 252 2946/2468
F: +44 (0)116 252 5799
E: criminology@le.ac.uk

APPENDIX E

Crime case scenario

The original PDF file created by the University of Hull (Higher Education Academy, JISC Open Educational Resources Programme, 2011) is contained in the addendum attached to the inside of the back cover of this thesis.

APPENDIX F

Crime case presentation

The Prezi presentation created by the researcher and used to present the fictional crime case scenario to the mock juries is available at: <https://prezi.com/>

Access to the presentation is subject to authorisation. To obtain authorisation, please email the researcher at: cd228@leicester.ac.uk

APPENDIX G

Prosecution and Defence closing arguments

PROSECUTION CASE SUMMARY

Ms Julie Ann Smith was raped and murdered on the night of 27 January 2017, between 19:00 and 22:00. There are no signs of forced entry, so we know that Julie Ann opened the door that night to someone she knew and was then raped and killed by someone she knew. It is the Prosecution's contention that that person was Mr Carl Henry Piggott, Julie Ann's neighbour, who – continuing his attempts to begin a sexual relationship with the victim – on the evening of the 27 January, raped and murdered Julie Ann.

How do we get to this conclusion? Let's reconstruct the facts:

- We know from Ms Emily Elliot, Julie Ann's flatmate, that there was no ongoing sexual relationship between Carl Piggott and Julie Ann and that, in fact, Mr Piggott had been "bothering" the victim for a matter of weeks. Julie Ann was annoyed by that behaviour, she said to Emily: 'that little creep from upstairs keeps on at me'.
- Now, there are no doubts that this interest was not reciprocated; Julie Ann would never have had sex with Carl Piggott, right? And yet Carl Piggott's semen was found at the crime scene in a condom (photo) recovered at the apartment.
- So, the only conclusion we can draw is that there was sexual intercourse that night, but it wasn't consensual. It was rape, as shown also by the bruising to the genital area of the deceased. After all, any hypothesis of consensual intercourse has to be excluded anyway, as Mr Piggott himself denied the occurrence of any sexual contact with the victim. Why would he have denied it, if it had been consensual?
- Let's also check the times (photo)... Julie Ann was killed between 7pm and 10pm. A neighbour, Ms Sarah Smith, told us that she saw Mr Piggott leaving

the building at 10pm that night and that she found it quite strange as she had not seen Mr Piggott leave at that time regularly... So, Carl Piggott did not go back to his apartment; he perhaps needed to take a walk, and he left at 10pm, after raping and killing Julie Ann. Now the defence will try to convince you that Mr Piggott leaving at that unusual time finds reason in the fact that he had just started a new job at a petrol garage. However, please bear in mind that, even if that was confirmed – which it wasn't, as the defendant results officially unemployed – that does not prove that Mr Piggott didn't commit the crimes: Julie Ann was killed between 7pm and 10pm – which means: he raped and killed her and then left at 10pm.

- And then we have the murder weapon (photo), a gun, that was also located at the crime scene and had Piggott's fingerprints on it. The defendant's fingerprints on the murder weapon... Does this need any more explanation?
- Now on the basis of what emerged during this trial, the defence will also try to draw your attention on some external, irrelevant factors – such as the potential involvement of Mr Frank Jones, Julie Ann's ex-boyfriend. But Frank says that he wasn't at Julie Ann's that night and, after all it isn't Frank's semen and fingerprints that were found at the crime scene.

Ladies and gentlemen of the jury, on the night of 27 January 2017, Mr Carl Henry Piggott went to Julie Ann Smith's flat with the intention to show once again his feelings for her and, when faced with another rejection, raped and killed her. There is a motive, there is the defendant's semen at the crime scene, there is the murder weapon with the defendant's fingerprints on it. There is no doubt he did it.

DEFENCE CASE SUMMARY

In response to the Prosecution's claims, the defence argues that Mr Piggott did not commit either of the crimes he has been accused of, and cannot be found guilty beyond reasonable doubt.

And here is why:

- First of all the presence of the defendant's DNA in the condom does not prove that he raped the victim, and even less that he killed her. The arguments presented by the prosecution do not prove that the sexual intercourse was not consensual – especially when considering that the victim was found fully clothed, which is highly inconsistent with the hypothesis of rape: what sex offender would dress their victim again after raping and killing her?! Therefore, if there was sexual intercourse, that must have been consensual – Mr Piggott's denial about any sexual relationship with the victim was obviously due to the fear of being accused of the murder, that's why he lied.
- Do you really think it's unbelievable that Julie Ann would have anything to do with 'that little creep from upstairs'? She was annoyed, perhaps disgusted, so she wouldn't get close to him. Would she accept to drink some wine with him if he brought her a bottle? Maybe not. Would she share a bottle with him... drinking from the same bottle?! Definitely not. Right? But what if I told you that there was an empty wine bottle (photo)... There were no glasses at the crime scene, and the DNA from both Julie Ann and Carl was found on the rim of the bottle.
- So why did Julie Ann lie to Emily about the nature of her relationship with Carl? Well, we don't know that. It might be that she felt ashamed or embarrassed as she was dating an older man, or perhaps it wasn't a serene or healthy relationship, but given the evidence we can't exclude a consensual sexual intercourse between the two that night.
- The prosecution highlighted that, according to a neighbour – Ms Smith – Mr Piggott left the building at 10pm and that this was unusual... Yet, this statement is compatible with the account of the defendant, who stated he left to go to work at 10pm. Was this unusual? Yes, because he had just started a new job a few days earlier – which also explains why he resulted officially unemployed.
- The prosecution also urged you not to consider factors that emerged during the trial and are irrelevant – such as the potential involvement of Frank Jones in these murders. But what if that weren't an irrelevant factor of this case? We

are here to find the truth, so let's look for it! Emily, the victim's flatmate, made mention of another important aspect of Julie Ann's life: she had split from her ex-boyfriend, Frank Jones, eight days before the murder, after a particularly nasty argument. It seems that there were some financial issues between the victim and Mr Jones. In the light of this, doesn't it seem interesting that Mr Jones' bank statement (photo) was also found at the crime scene? Doesn't it suggest that perhaps Julie Ann's ex-boyfriend was in the apartment that night? A nasty argument over financial problems... that Mr Jones might have wanted to try and solve that night... and that instead turned into tragedy.

- And somebody shot Julie Ann. The prosecution mentioned a gun as murder weapon. And that's right. However there was another gun (photo) at the crime scene and it has yet to be established which one was actually the murder weapon, as ballistics were inconclusive.
- Also, the fingerprints found on the first gun are of a thumb and forefinger (of the left hand) and they were found on the barrel of the gun (not the trigger). This means that they don't prove that the person whose fingerprints are on it held the gun to shoot.

Ladies and gentlemen of the jury, we all wish that this didn't happen and we are all here to find out the truth. The decision is now yours to make. So please consider the facts of this case and render your verdict.

APPENDIX H

Demographic questionnaire and notes

JUROR n. ____

- What is your age?
 - ☐ 25 or under
 - ☐ 26-40
 - ☐ 41-55
 - ☐ 56 or older

- What is your gender?
 - ☐ Male
 - ☐ Female

- What is your nationality?

- What is your occupation?

- Have you ever served on a jury?
 - ☐ Yes
 - ☐ No

YOUR NOTES ABOUT THE CASE

APPENDIX I

Pre-/Post-deliberation questionnaires

Pre-deliberation questionnaire

JUROR n. __

1) After the presentation of this (fictional) crime case, do you think the defendant is:

- ☐ Guilty on both counts
- ☐ Not guilty on both counts
- ☐ Guilty of rape; not guilty of murder
- ☐ Not guilty of rape; guilty of murder

2) How confident are you in your decision?

1 Extremely unsure	2 Very unsure	3 Not sure	4 Very confident	5 Extremely confident

(Please tick as appropriate)

Post-deliberation questionnaire (British condition)

JUROR n. ____

- 1) You and the other members of the (mock) jury have just reached a verdict on the case presented. How do you feel about the fairness of the decision taken?

The decision was:

1	2	3	4	5
Extremely unfair	Very unfair	Not sure how fair	Very fair	Extremely fair

(Please tick as appropriate)

- 2) Does the group decision reflect your individual opinion about the guilt/innocence of the defendant?

☐ Yes ☐ No

- 3) Regardless of the group decision, do you individually think the defendant is:

- ☐ Guilty on both counts
- ☐ Not guilty on both counts
- ☐ Guilty of rape; not guilty of murder
- ☐ Not guilty of rape; guilty of murder

4) How confident are you in your conviction/opinion?

1 Extremely unsure	2 Very unsure	3 Not sure	4 Very confident	5 Extremely confident

(Please tick as appropriate)

5) Were there members of the panel who tended to lead or prevail in the discussions?

☐ Yes ☐ No

6) During the deliberation, did you think there was need for someone to direct the discussion?

☐ Yes ☐ No

7) Do you think the presence of professionals (e.g. judges) on the jury panel might have been of help for you?

☐ Yes ☐ No

8) After the discussion with the other members of the mock jury, do you think you could provide a reason for your choice (guilty/not guilty)?

☐ Yes ☐ No

(If Yes) What is/are the reason/s for your choice (guilty/not guilty)?

(Please briefly explain below)

--

Post-deliberation questionnaire (Italian condition – translated from Italian)

JUROR n. ____

- 1) You and the other members of the (mock) jury have just reached a verdict on the case presented. How do you feel about the fairness of the decision taken?

The decision was:

1	2	3	4	5
Extremely unfair	Very unfair	Not sure how fair	Very fair	Extremely fair

(Please tick as appropriate)

- 2) Does the group decision reflect your individual opinion about the guilt/innocence of the defendant?

☐ Yes ☐ No

- 3) Regardless of the group decision, do you individually think the defendant is:

- ☐ Guilty on both counts
- ☐ Not guilty on both counts
- ☐ Guilty of rape; not guilty of murder
- ☐ Not guilty of rape; guilty of murder

4) How confident are you in your conviction/opinion?

1	2	3	4	5
Extremely unsure	Very unsure	Not sure	Very confident	Extremely confident

(Please tick as appropriate)

5) Were there members of the panel who tended to lead or prevail in the discussions?

☐ Yes ☐ No

6) Did you find it useful that, during the deliberation, there was someone there (namely, the “judge”) directing the discussion?

☐ Yes ☐ No

7) After the discussion with the other members of the mock jury, do you think you could provide a reason for your choice (guilty/not guilty)?

☐ Yes ☐ No

(If Yes) What is/are the reason/s for your choice (guilty/not guilty)?

(Please briefly explain below)

--

BIBLIOGRAPHY

- Ainsworth, P.B. (1998) *Psychology, law, and eyewitness testimony*. Chichester: Wiley.
- Aitken, C.G. and Stoney, D.A. (1991) *The use of statistics in forensic science*. CRC Press.
- Anderson, N.H. (1968) 'A simple model for information integration', *Theories of cognitive consistency: A sourcebook*, pp. 731-743.
- Anderson, N.H. (1981) 'Foundations of information integration theory', New York: Academic Press.
- Anderson, T., Schum, D. and Twining, W. (2005) *Analysis of evidence*. Cambridge University Press.
- Apa, S. and Cantarini, L., (2011) *Il ruolo della giuria in Italia e negli USA*. PhD Thesis, Sapienza – Università di Roma.
- Arkes, H.R. and Mellers, B.A. (2002) 'Do juries meet our expectations?', *Law and human behavior*, 26(6), pp. 625-639.
- Asch, S. E. (1951) 'Effects of group pressure upon the modification and distortion of judgment', in H. Guetzkow (ed.) *Groups, leadership and men*. Pittsburgh, PA: Carnegie Press.
- Bankert, E.A. and Amdur, R.J. (2006) *Institutional review board: Management and function*. Sudbury, Massachusetts: Jones and Bartlett Publishers.
- Baron, J. (2008) *Thinking and deciding*. New York: Cambridge University Press.
- Bailey, J.E. and Pearson, S.W. (1983) 'Development of a tool for measuring and analyzing computer user satisfaction', *Management science*, 29(5), pp. 530-545.
- Beresford, B. and Sloper, P. (2008) *Understanding the dynamics of decision-making and choice: A scoping study of key psychological theories to inform the design and analysis of the panel study*. Social Policy Research Unit, University of York York.
- Bernard, H.R. and Ryan, G.W. (2010) *Analyzing qualitative data: systematic approaches*. Los Angeles; London: SAGE.
- Berti, C. (2002) *Psicologia sociale della giustizia*. Bologna: Il Mulino.
- Bex, F., Prakken, H., and Verheij, B. (2006) 'Anchored Narratives in Reasoning about Evidence', *Proceedings of the 2006 conference on Legal Knowledge and Information Systems: The Nineteenth Annual Conference, Amsterdam*: IOS Press, 11-20.

- Bicchieri, C. (2013) 'Words and Deeds: a Focus Theory of Norms', in Nida-Rümelin, J. and Spohn, W. (eds.) *Rationality, Rules, and Structure*. Dordrecht: Springer, pp. 153-184.
- Billig, M. (1986) *Thinking and arguing*. Loughborough: Loughborough University of Technology.
- Binchi, C., 2014 'Repressione del dissenso e giustizia penale: la Corte d'Assise di Bologna', *Percorsi Storici: Rivista di Storia Contemporanea*.
- Bird, C.M. (2005) 'How I Stopped Dreading and Learned to Love Transcription', *Qualitative Inquiry*, 11(2), pp. 226-248.
- Bjerregaard, B.E., Smith, M.D., Cochran, J.K. and Fogel, S.J. (2017) 'A further examination of the liberation hypothesis in capital murder trials', *Crime & Delinquency*, 63(8), pp. 1017-1038.
- Blackledge, R. and Vincenti, M. (1994) 'Identification of polydimethylsiloxane lubricant traces from latex condoms in cases of sexual assault', *Journal of the Forensic Science Society*, 34(4), pp. 245-256.
- Blackledge, R.D. (1996) 'Condom trace evidence: a new factor in sexual assault investigations', *FBI L. Enforcement Bull*, 65(5), pp. 12-17.
- Blackmun, H.A. (1994) *Victor v. Nebraska*. 511 U.S. 1, 29 edn.
- Blaikie, N.W. (1991) 'A critique of the use of triangulation in social research', *Quality and quantity*, 25(2), pp. 115-136.
- Bleich, E. and Pekkanen, R. (2013) 'How to report interview data', in Mosley, L. (ed.) *Interview Research in Political Science*. New York: Cornell University Press, pp. 84-105.
- Bond, R. (2005) 'Group Size and Conformity', *Group Processes & Intergroup Relations*, 8(4), pp. 331-354.
- Bornstein, B.H. (1999) 'The Ecological Validity of Jury Simulations: Is the Jury Still Out?', *Law and Human Behavior*, 23(1), pp. 75-91.
- Bornstein, B.H., Golding, J.M., Neuschatz, J., Kimbrough, C., Reed, K., Magyarics, C. and Luecht, K. (2017) 'Mock juror sampling issues in jury simulation research: A meta-analysis.', *Law and human behavior*, 41(1), pp. 13-28.
- Bradshaw, R., Wolstenholme, R., Ferguson, L.S., Sammon, C., Mader, K., Claude, E., Blackledge, R.D., Clench, M.R. and Francese, S. (2013) 'Spectroscopic imaging based approach for condom identification in condom contaminated fingerprints', *Analyst*, 138(9), pp. 2546-2557.

- Braun, V. and Clarke, V., 2006. 'Using thematic analysis in psychology', *Qualitative research in psychology*, 3(2), pp. 77-101.
- Bray, R.M. and Kerr, N.L. (1979) 'Use of the simulation method in the study of jury behavior: Some methodological considerations', *Law and Human Behavior*, 3, pp. 107-119.
- Bray, R.M. and Kerr, N.L. (1982) 'Methodological considerations in the study of the psychology of the courtroom', in Kerr, N. and Bray, R. (eds.) *The psychology of the courtroom*. London: Academic Press, pp. 287-324.
- Brooks, T. (2004) 'The right to trial by jury', *Journal of Applied Philosophy*, 21(2), pp. 197-212.
- Brooks, J., McCluskey, S., Turley, E. and King, N. (2015) 'The Utility of Template Analysis in Qualitative Psychology Research', *Qualitative Research in Psychology*, 12(2), pp. 202-222.
- Bryman, A. (2012) *Social research methods*. Oxford: Oxford University Press.
- Burnett, A. and Badzinski, D.M. (2000) 'An exploratory study of argument in the jury decision-making process', *Communication Quarterly*, 48(4), pp. 380-396.
- Carlson, K.A. and Russo, J.E. (2001) 'Biased Interpretation of Evidence by Mock Jurors', *Journal of Experimental Psychology: Applied*, 7(2), pp. 91-103.
- Carneiro, J., Santos, R., Marreiros, G. and Novais, P. (2014) 'Understanding decision quality through satisfaction', *International Conference on Practical Applications of Agents and Multi-Agent Systems* Springer, pp. 368-377.
- CPP (2013), *Codice di procedura penale italiano, Libro II, Le corti di assise*. Milano: L. di Giacomo Pirola.
- Champagne, A. and Nagel, S. (1982) 'The psychology of judging', *The psychology of the courtroom*, pp. 221-256.
- Chankong, V. and Haimes, Y.Y. (2008) *Multiobjective decision making: theory and methodology*. Courier Dover Publications.
- Charron, A. and Woodhams, J. (2010) 'A qualitative analysis of mock jurors' deliberations of linkage analysis evidence', *Journal of Investigative Psychology and Offender Profiling*, 7(2), pp. 165-183.
- Charrow, R.P. and Charrow, V.R. (1979) 'Making legal language understandable: A psycholinguistic study of jury instructions', *Columbia law review*, 79(7), pp. 1306-1374.

- Chemers, M.M. (2000) 'Leadership research and theory: A functional integration', *Group Dynamics: Theory, research, and practice*, 4(1), pp. 27-43.
- Chesterman, M.R., Chan, J.B. and Hampton, S. (2001) *Managing prejudicial publicity: an empirical study of criminal jury trials in New South Wales*. Justice Research Centre, Law and Justice Foundation of NSW.
- Chilton, E. and Henley, P. (1996) 'Improving the jury system, jury instructions: Helping jurors understand the evidence and the law', *University of California, Hastings College of the Law, Public Law Research Institute*.
- Chong, E. and Wolf, H. (2010) 'Factors influencing followers' perception of organisational leaders', *Leadership & Organization Development Journal*, 31(5), pp. 402-419.
- Christiaansen, R.E., Sweeney, J.D. and Ochalek, K. (1983) 'Influencing eyewitness descriptions', *Law and Human Behavior*, 7(1), pp. 59-65.
- Clark, N.K., Stephenson, G. and Rutter, D. (1986) 'Memory for a complex social discourse: The analysis and prediction of individual and group recall', *Journal of Memory and Language*, 25(3), pp. 295-313.
- Cleary, M., Walter, G. and Matheson, S. (2008) 'The challenge of optimising research participation: paying participants in mental health settings', *Acta neuropsychiatrica*, 20(6), pp. 286-290.
- Cole, S.A. and Dioso-Villa, R. (2007) 'CSI and its effects: media, juries, and the burden of proof.(The CSI Effect: The True Effect of Crime Scene Television on the Justice System)', *New England Law Review*, 41(3), pp. 435-757.
- Cole, S.A. and Dioso-Villa, R. (2009) 'Investigating the 'CSI effect' effect: media and litigation crisis in criminal law. ('CSI: Crime Scene Investigation' television program) (Symposium: Media, Justice, and the Law)', *Stanford law review*, 61(6), pp. 1335-1339.
- Cooley, C.M. (2007) 'The CSI effect: its impact and potential concerns.(Crime Scene Investigation)(The CSI Effect: The True Effect of Crime Scene Television on the Justice System)', *New England Law Review*, 41(3), pp. 471-501.
- Corallo, G., J. Sackur, S. Dehaene, and M. Sigman (2008) 'Limits on Introspection: Distorted Subjective Time During the Dual-Task Bottleneck', *Psychological Science*, 19 (11), 1110-1117.
- Creswell, J.W. (2012) *Qualitative inquiry and research design: choosing among five approaches*. Thousand Oaks, Calif.: Sage Publications.
- Creswell, J.W. and Plano Clark, V.L. (2007) *Designing and conducting mixed methods research*. Thousand Oaks, Calif.; London: Sage Publications.

- Crotty, M. (1998) *The foundations of social research: meaning and perspective in the research process*. London: Sage Publications.
- Crown Court Compendium (2017) Available at: <https://www.judiciary.uk/publications/crown-court-bench-book-directing-the-jury-2/> (Accessed: 06/02/2018).
- da Luz, C.M. and Weckerly, P.C. (1993) 'The Texas Condom-Rape Case: Caution Construed as Consent', *UCLA Women's Law Journal*, 3(0), pp. 95-104.
- Davis, J.H., Kerr, N.L., Atkin, R.S., Holt, R. and Meek, D. (1975) 'The decision processes of 6- and 12- person mock juries assigned unanimous and two-thirds majority rules', *Journal of personality and social psychology*, 32(1), pp. 1-14.
- De Dreu, C.K., Nijstad, B.A. and van Knippenberg, D. (2008) 'Motivated information processing in group judgment and decision making', *Personality and social psychology review*, 12(1), pp. 22-49.
- De La Fuente, L., De La Fuente, E Inmaculada and García, J. (2003) 'Effects of pretrial juror bias, strength of evidence and deliberation process on juror decisions: New validity evidence of the Juror Bias Scale scores', *Psychology, Crime & Law*, 9(2), pp. 197-209.
- Demaine, L.J. (2008) 'In search of an anti-elephant: Confronting the human inability to forget inadmissible evidence', *Geo. Mason L. Rev.*, 16, pp. 99-101.
- DeMatteo, D. and Anumba, N. (2009) 'The validity of jury decision making research', in Lieberman, J. and Krauss, D. (eds.) *Jury psychology: Social aspects of trial processes*. Surrey: Ashgate Publishing Limited, pp. 1-23.
- Denzin, N.K. and Lincoln, Y.S. (1994) *Handbook of qualitative research*. Thousand Oaks, Calif.; London: Sage Publications.
- Deutsch, M. and Gerard, H.B. (1955) 'A study of normative and informational social influences upon individual judgment', *The journal of abnormal and social psychology*, 51(3), pp. 629-636.
- Devine, D.J. (2012) *Jury decision making: The state of the science*. NYU Press.
- Devine, D.J., Clayton, L.D., Dunford, B.B., Seying, R. and Pryce, J. (2001) 'JURY DECISION MAKING 45 Years of Empirical Research on Deliberating Groups', *Psychology, Public Policy, and Law*, 7(3), pp. 622-727.
- Devine, D.J., Buddenbaum, J., Houp, S., Studebaker, N. and Stolle, D.P. (2009) 'Strength of evidence, extraevidentiary influence, and the liberation hypothesis: Data from the field', *Law and Human Behavior*, 33(2), pp. 136-148.

- Devlin, P.A.D., Baron (1965) *The enforcement of morals*. London: Oxford University Press.
- Dhami, M.K., Lundrigan, S. and Mueller-Johnson, K. (2015) 'Instructions on Reasonable Doubt: Defining the Standard of Proof and the Juror's Task', *Psych. Pub. Pol. and L.*, 21, pp. 169-452.
- Di Majo, L., 2014 'Errori giudiziari e giuria: le garanzie di imparzialità e indipendenza dei giudici laici tra dato normativo e prassi', *Diritto.it*, Available at: <https://www.diritto.it/> (Accessed: 26 October 2015).
- Diamond, S.S. (1997) 'Illuminations and Shadows from Jury Simulations', *Law and human behavior*, 21(5), pp. 561-571.
- Diamond, S.S., Casper, J.D. and Ostergren, L. (1989) 'Blindfolding the jury', *Law and contemporary problems*, 52(4), pp. 247-267.
- Dietrich, F. and List, C. (2013) 'A Reason-Based Theory of Rational Choice', *Noûs*, 47(1), pp. 104-134.
- DiFonzo, J.H. (2005) 'The Crimes of Crime Labs', *Hofstra Law Review*, 34(1), pp. 1-11.
- DiFonzo, J.H. and Stern, R.C. (2006) 'Devil in a white coat: the temptation of forensic evidence in the age of CSI.(Crime Scene Investigation)(The CSI Effect: The True Effect of Crime Scene Television on the Justice System)', *New England Law Review*, 41(3), pp. 503-757.
- Drapeau, M. (2002) 'Subjectivity in research: Why not? But...', *The Qualitative Report*, 7(3), pp. 1-15.
- Draper, J. (2004) 'The relationship between research question and research design', in Crookes, Patrick A. and Davies, Sue (eds.) *Research into Practice: Essential Skills for Reading and Applying Research in Nursing and Health Care*, Edinburgh: Bailliere Tindall, pp. 69–84.
- Drever, E. (1995) *Using Semi-Structured Interviews in Small-Scale Research. A Teacher's Guide*. Glasgow: SCRE Centre, University of Glasgow.
- Drisko, J.W. (1997) 'Strengthening qualitative studies and reports: Standards to promote academic integrity', *Journal of Social Work Education*, 33 (1), pp. 185-197.
- Driscoll, D.L., Appiah-Yeboah, A., Salib, P. and Rupert, D.J. (2007) 'Merging qualitative and quantitative data in mixed methods research: How to and why not', *Ecological and Environmental Anthropology (University of Georgia)*, 3(1), pp. 18-28.

- Durant, W. (2011) *The Life of Greece: The Story of Civilization*. Simon and Schuster.
- Dutwin, D. (2003) 'The character of deliberation: Equality, argument, and the formation of public opinion', *International Journal of Public Opinion Research*, 15(3), pp.239-264.
- Eden, D. and Leviatan, U. (1975) 'Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales', *Journal of Applied Psychology*, 60(6), pp. 736-741.
- Ellison, L. and Munro, V.E. (2009) 'Reacting to Rape: Exploring Mock Jurors' Assessments of Complainant Credibility', *British Journal of Criminology*, 49(2), pp. 202-219.
- Ellison, L. and Munro, V.E. (2010a) 'Getting to (not) guilty: examining jurors' deliberative processes in, and beyond, the context of a mock rape trial', *Legal Studies*, 30(1), pp. 74-97.
- Ellison, L. and Munro, V.E. (2010b) 'A stranger in the bushes, or an elephant in the room? Critical reflections upon received rape myth wisdom in the context of a mock jury study', *New Criminal Law Review: In International and Interdisciplinary Journal*, 13(4), pp. 781-801.
- Ellison, L. and Munro, V.E. (2013) 'Better the Devil You Know? 'Real Rape' Stereotypes and the Relevance of a Previous Relationship in (Mock) Juror Deliberations', *The International Journal of Evidence & Proof*, 17(4), pp. 299-322.
- Ellison, L. and Munro, V.E. (2015) 'Telling tales': exploring narratives of life and law within the (mock) jury room', *Legal studies*, 35(2), pp.201-225.
- Ellsworth, P.C. (1989) 'Are twelve heads better than one?', *Law and contemporary problems*, 52(4), pp. 205-224.
- Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K. and Kyngäs, H. (2014) 'Qualitative content analysis: A focus on trustworthiness', *Sage Open*, 4(1), pp. 1-10.
- Elwork, A., Sales, B.D. and Alfini, J.J. (1977) 'Juridic decisions: In ignorance of the law or in light of it?', *Law and human behavior*, 1(2), pp. 163-189.
- Emrich, C.G. (1999) 'Context effects in leadership perception', *Personality and Social Psychology Bulletin*, 25(8), pp. 991-1006.
- Fadalti, L. (2008) *La testimonianza nel giudizio penale*. Milano: Giuffrè Editore.
- Faigman, D. and Baglioni, A., (1988). 'Bayes' theorem in the trial process'. *Law and Human Behavior*, 12, pp. 1-17.

- Fiedler, F.E. (2006) 'The contingency model: A theory of leadership effectiveness' *Small groups: Key readings*, 12(4), pp.369-382.
- Filmer, P. (1972) *New directions in sociological theory*. London: Collier-Macmillan
- Fisher, A. (1974) *Small Group Decision Making: Communication and the Group Process*. London: McGraw-Hill.
- Fisher, M.A. (2000) 'Going for the Blue Ribbon: The Legality of Expert Juries in Patent Litigation', *Colum. Sci. & Tech. L. Rev.*, 2, pp. 2-3.
- Forman, J. and Damschroder, L. (2007) 'Qualitative content analysis' in *Empirical methods for bioethics: A primer*. Emerald Group Publishing Limited, pp. 39-62.
- Forsyth, W. (1852) *History of trial by jury*. The Lawbook Exchange, Ltd.: England, United Kingdom.
- Foss, R.D. (1976) 'Group decision processes in the simulated trial jury', *Sociometry*, 39(4), pp. 305-316.
- Frank, J. (1973) *Courts on trial: Myth and reality in American justice*. Princeton University Press.
- Frey, B.S. (2005) "'Just forget it.'" Memory distortions as bounded rationality', *Mind & Society*, 4(1), pp. 13-25.
- Frier, B.W. (1985) *The rise of the Roman jurists: studies in Cicero's Pro Caecina*. Guildford; Princeton, N.J.: Princeton University Press.
- Fujita, M. and Hotta, S. (2010) 'The impact of differential information between lay participants and professional judges on deliberative decision-making', *International Journal of Law, Crime and Justice*, 38(4), pp.216-235.
- García-Bajos, E. and Migueles, M. (2003) 'False memories for script actions in a mugging account', *European Journal of Cognitive Psychology*, 15(2), pp. 195-208.
- Gee, J.P. (2011) *An introduction to discourse analysis theory and method*. New York: Routledge.
- Giessner, S.R., van Knippenberg, D.L. and Sleebos, E. (2008) 'License to fail? How leader group prototypicality moderates the effects of leader performance on perceptions of leadership effectiveness', 20(3), pp. 434-451.
- Gill, P., Stewart, K., Treasure, E. and Chadwick, B. (2008) 'Methods of data collection in qualitative research: interviews and focus groups', *British dental journal*, 204(6), pp. 291-295.
- Gini, A. and Green, R.M. (2013) *What Is Leadership?* Wiley-Blackwell: Oxford, UK.

- Gleisser, M. (1968) *Juries and justice*. Barnes; Yoseloff.
- Glöckner, A. and Betsch, T. (2008) 'Multiple-reason decision making based on automatic processing', *Journal of experimental psychology: Learning, memory, and cognition*, 34(5), pp. 1055-1075.
- Glöckner, A. and Engel, C. (2013) 'Can we trust intuitive jurors? Standards of proof and the probative value of evidence in coherence-based reasoning', *Journal of Empirical Legal Studies*, 10(2), pp. 230-252.
- Goldstein, A.S. (1993) 'Jury Secrecy and the Media: The Problem of Postverdict Interviews', *University of Illinois Law Review*, 1993(2), pp. 295-314.
- Grady, C. (2005) 'Payment of clinical research subjects', *The Journal of clinical investigation*, 115(7), pp. 1681-1687.
- Greene, E., Chopra, S.R., Kovera, M.B., Penrod, S.D., Rose, V.G., Schuller, R. and Studebaker, C.A. (2002) 'Jurors and Juries', in Ogloff, J.R.P. (ed.) *Taking psychology and law into the twenty-first century*. New York: Kluwer Academic/Plenum., pp. 225-284.
- Groot, R.D. (1988) 'The Early-Thirteenth-Century Criminal Jury', in Cockburn, J.S. and Green, T.A. (eds.) *Twelve Good Men and True: The Criminal Trial Jury in England, 1200-1800*. Princeton, New Jersey: Princeton University Press, pp. 3-35.
- Groscup, J. and Tallon, J. (2009) 'Theoretical models of jury decision-making', *Psychology, crime and law series. Jury psychology: Social aspects of trial processes: Psychology in the courtroom*, 1, pp. 41-65.
- Grove, T. (2012) *The Juryman's Tale*. Bloomsbury Publishing.
- Grubin, D., Kelly, P. and Brunson, C. (2001) *Linking serious sexual assaults through behaviour*. Home Office, Research, Development and Statistics Directorate London.
- Gulotta, G. (1987) *Trattato di psicologia giudiziaria nel sistema penale*. Milano: Giuffrè.
- Gunby, C., Carline, A. and Beynon, C. (2013) 'Regretting it After? Focus Group Perspectives on Alcohol Consumption, Nonconsensual Sex and False Allegations of Rape', *Social & Legal Studies*, 22(1), pp. 87-106.
- Hall, D.F., Loftus, E.F. and Tausignant, J.P. (1984) 'Postevent information and changes in recollection for a natural event', *Eyewitness testimony: Psychological perspectives*, pp. 124-141.
- Hannigan, S.L. and Tippens Reinitz, M. (2001) 'A demonstration and comparison of two types of inference-based memory errors', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(4), pp. 931-940.

- Hans, V.P. (2008) 'Jury Systems Around the World', *Annual Review of Law and Social Science*, 4(1), pp. 275-297.
- Hansen, M.H. (2003) *La democrazia ateniese nel IV sec. a.C.* Milano: LED.
- Harrison, C. (2018) 'Leadership Theory and Research', in *Leadership Theory and Research: A Critical Approach to New and Existing Paradigms*. London: Palgrave Macmillan, pp. 15-32.
- Hastie, R. (1994) *Inside the juror: The Psychology of Juror Decision Making*. Cambridge: Cambridge University Press.
- Hastie, R., Penrod, S. and Pennington, N. (1983) *Inside the jury*. The Lawbook Exchange, Ltd.
- Hastie, R., Penrod, S. and Pennington, N. (2002) *Inside the jury*. New Jersey: The Lawbook Exchange, LTD.
- Hastie, R., Penrod, S. and Pennington, N. (2013) *Inside the jury*. The Lawbook Exchange, Ltd.
- Hastie, R. and Dawes, R.M. (2010) *Rational choice in an uncertain world: an introduction to judgment and decision making*. London: Sage.
- Heller, K.J. (2006) 'The cognitive psychology of circumstantial evidence', *Michigan Law Review*, pp. 241-305.
- Hendry, S.H., Shaffer, D.R. and Peacock, D. (1989) 'On testifying in one's own behalf: Interactive effects of evidential strength and defendant's testimonial demeanor on mock jurors' decisions', *Journal of Applied Psychology*, 74(4), pp. 539-545.
- Higgins, E.T. (2000) 'Making a good decision: value from fit', *American psychologist*, 55(11), pp. 1217-1230.
- Higher Education Academy, JISC Open Educational Resources Programme, 2011.
- Hinsz, V.B. (1990) 'Cognitive and consensus processes in group recognition memory performance.', *Journal of personality and social psychology*, 59(4), pp. 705-718.
- Holyoak, K.J. and Simon, D. (1999) 'Bidirectional reasoning in decision making by constraint satisfaction', *Journal of Experimental Psychology: General*, 128(1), pp. 3-31.
- Hosch, H.M., Beck, E.L. and McIntyre, P. (1980) 'Influence of Expert Testimony Regarding Eyewitness Accuracy on Jury Decisions', *Law and human behavior*, 4(4), pp. 287-296.

- Hostettler, J. (2004) *The criminal jury old and new: jury power from early times to the present day*. Winchester: Waterside Press.
- Houck, M.M. (2006) 'CSI: reality', *Scientific American*, 295(1), pp. 84-89.
- House, R.J. and Mitchell, T.R. (1975) *Path-goal theory of leadership*. Washington University: Seattle Dept Of Psychology.
- Howe, M.L. and Knott, L.M. (2015) 'The fallibility of memory in judicial processes: Lessons from the past and their modern consequences', *Memory*, 23(5), pp. 633-656.
- Hsieh, H. and Shannon, S.E. (2005) 'Three Approaches to Qualitative Content Analysis', *Qualitative health research*, 15(9), pp. 1277-1288.
- Imwinkelried, E.J. (1983) 'The standard for admitting scientific evidence: a critique from the perspective of juror psychology', *Villanova Law Review*, 28(3), pp. 554-571.
- Ivković, S.K. and Hans, V.P. (2003) 'Jurors' evaluations of expert testimony: Judging the messenger and the message', *Law & Social Inquiry*, 28(2), pp. 441-482.
- Ivković, S.K. (2003) 'An Inside View: Professional Judges' and Lay Judges' Support for Mixed Tribunals', *Law & Policy*, 25(2), pp. 93-122.
- Jenness, A. (1932) 'The role of discussion in changing opinion regarding a matter of fact', *The Journal of Abnormal and Social Psychology*, 27(3), pp. 279-296.
- Jick, T.D. (1979) 'Mixing Qualitative And Quantitative Methods: Triangulation In Action', *Administrative Science Quarterly*, 24, pp. 602-611.
- Johnson, R.B. and Turner, L.S. (2003) 'Data collection strategies in mixed methods research.', in Tashakkori, A. and Teddlie, C. (eds.) *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage, pp. 297-319.
- Jupp, D.V.R. (2012) *Methods of Criminological Research*. Hoboken: Taylor and Francis.
- Kahneman, D. and Tversky, A. (1984) 'Choices, values, and frames.', *American psychologist*, 39(4), pp. 341-350.
- Kalven Jr, H. and Zeisel, H. (1966) *The American Jury*. Chicago: University of Chicago.
- Kaplan, M.F. and Martín, A.M. (2006) 'Introduction and Overview', in Kaplan, M.F. and Martín, A.M. (eds.) *Understanding World Jury Systems Through Social Psychological Research*. New York and Hove: Psychology Press, pp. 1-8.
- Kaplan, M.F., Martín, A.M. and Hertel, J. (2006) 'Issues and Prospects in European Juries: An Overview', in Kaplan, M.F. and Martín, A.M. (eds.) *Understanding world*

- jury systems through social psychological research*. New York and Hove: Psychology Press, pp. 111-124.
- Kaplan, M.F. and Miller, L.E. (1978) 'Reducing the effects of juror bias', *Journal of personality and social psychology*, 36(12), pp. 1443-1455.
- Kassin, S.M. and Kiechel, K.L. (1996) 'The social psychology of false confessions: Compliance, internalization, and confabulation', *Psychological science*, 7(3), pp. 125-128.
- Kassin, S.M. and Wrightsman, L.S. (1988) *The American jury on trial: psychological perspectives*. New York: Hemisphere Publishing.
- Katzenbach, J. R. and Smith, D.K. (1993) *The Wisdom of Teams: Creating the High-performance Organisation*. Harvard Business School: Boston.
- Kaye, D.H. and Koehler, J.J. (1991) 'Can jurors understand probabilistic evidence?', *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 154(1), pp. 75-81.
- Kelman, H.C. (1958) 'Compliance, identification, and internalization three processes of attitude change', *Journal of Conflict Resolution*, 2(1), pp. 51-60.
- Kessler, J.B. (1975) 'The social psychology of jury deliberations', *The jury system in America: A critical overview*. Beverly Hills: Sage Publications.
- Kılıç, A., Criss, A.H., Malmberg, K.J. and Shiffrin, R.M. (2017) 'Models that allow us to perceive the world more accurately also allow us to remember past events more accurately via differentiation', *Cognitive psychology*, 92, pp. 65-86.
- King, A. (1993) 'From Sage on the Stage to Guide on the Side', *College Teaching*, 41(1), pp. 30-35.
- King, N., 2012. 'Doing Template Analysis', in Symon, G. and Cassell, C. (eds), *Qualitative organizational research: Core methods and current challenges*. London: Sage, pp. 426-450.
- Kirgis, P.F. (2002) 'The Problem of the Expert Juror', *Temple Law Review*, 75(3), pp. 493-537.
- Knight, F.H. (2012) *Risk, uncertainty and profit*. Boston, New York: Houghton Mifflin Co.
- Kovera, M.B., McAuliff, B.D. and Hebert, K.S. (1999) 'Reasoning about scientific evidence: Effects of juror gender and evidence quality on juror decisions in a hostile work environment case', *Journal of Applied Psychology*, 84(3), pp. 362-375.
- Krueger, R.A. (1994) *Focus Group: Practical Guide for Applied Research*. London: Sage.

- Kuhn, D., Weinstock, M. and Flaton, R. (1994) 'How well do jurors reason? Competence dimensions of individual variation in a juror reasoning task', *Psychological Science*, 5(5), pp. 289-296.
- Langbein, J.H. (1981) 'Mixed Court and Jury Court: Could the Continental Alternative Fill the American Need?', *American Bar Foundation Research Journal*, 6(1), pp. 195-219.
- Langerock, H. (1915) 'Professionalism: A Study in Professional Deformation', *American Journal of Sociology*, 21(1), pp. 30-44.
- Lanza, L. (1994) *Gli omicidi in famiglia: le dinamiche della decisione in Corte d'assise con l'analisi di trenta casi di omicidio domestico*. Milano: Giuffrè.
- Lapadat, J.C. and Lindsay, A.C. (1999) 'Transcription in Research and Practice: From Standardization of Technique to Interpretive Positionings', *Qualitative Inquiry*, 5(1), pp. 64-86.
- Laudan, L. (2003) 'Is Reasonable Doubt Reasonable?', *Legal Theory*, 9(4), pp. 295-331.
- Laughlin, P.R. (2011) 'Social choice theory, social decision scheme theory, and group decision-making', *Group Processes & Intergroup Relations*, 14(1), pp. 63-79.
- Leedy, P.D. and Ormrod, J.E. (2013) *Practical Research: Planning and Design*. Boston, Mass: Pearson Education International.
- Lempert, R. (1991) 'Telling tales in court: Trial procedure and the story model', *Cardozo L. Rev.*, 13(2-3), pp. 559-573.
- Lloyd-Bostock, S.M. and British Psychological Society (1988) *Law in practice: applications of psychology to legal decision making and legal skills*. London: British Psychological Society and Routledge.
- Loftus, E.F., (1975) 'Leading questions and the eyewitness report', *Cognitive Psychology*, 7(4), pp. 560-572.
- Loftus, E.F. (1979) 'The malleability of human memory: Information introduced after we view an incident can transform memory', *American Scientist*, 67(3), pp. 312-320.
- Loftus, E.F. (1980) 'Impact of expert psychological testimony on the unreliability of eyewitness identification' *Journal of Applied Psychology*, 65(1), p.9-15.
- Loftus, E.F. (1996a) *Eyewitness testimony*. Cambridge, Mass: Harvard University Press.
- Loftus, E. (1996b) 'Memory distortion and false memory creation', *Bulletin of the American Academy of Psychiatry and Law*, 24(3), pp. 281-295.

- Loftus, E.F. (2000) 'Remembering what never happened', in Tulving, E. (ed) *Memory, consciousness, and the brain: The Tallinn Conference*. Psychology Press/Taylor & Francis Philadelphia, pp. 106-118.
- Loftus, E.F. and Loftus, G.R. (1980) 'On the permanence of stored information in the human brain', *American Psychologist*, 35(5), pp. 409-420.
- London, K. and Nunez, N. (2000) 'The effect of jury deliberations on jurors' propensity to disregard inadmissible evidence', *Journal of Applied Psychology*, 85(6), pp. 932-939.
- Longhurst, R. (2010) 'Semi-structured interviews and focus groups', in Clifford, N., French, S. and Valentine G. (eds) *Key Methods in Geography*. London: Sage, pp. 103-115.
- Lynch, J.F. (2013) 'Aligning Sampling Strategies with Analytic Goals', in Mosley, L. (ed.) *Interview Research in Political Science*. New York: Cornell University Press, pp. 31-44.
- Lyon, M.R. (2004) 'No means no?: Withdrawal of consent during intercourse and the continuing evolution of the definition of rape', *The Journal of Criminal Law and Criminology (1973-)*, 95(1), pp. 277-314.
- Ma, Y. (1998) 'Lay Participation in Criminal Trials: A Comparative Perspective', *International Criminal Justice Review*, 8 (1), pp. 74-94.
- MacCoun, R.J. (1989) 'Experimental research on jury decision-making', *Science (New York, N.Y.)*, 244(4908), pp. 1046-1050.
- MacCoun, R.J. and Kerr, N.L. (1988) 'Asymmetric influence in mock jury deliberation: Jurors' bias for leniency', *Journal of personality and social psychology*, 54(1), pp. 21-33.
- Maki, R.H., Weigold, A. and Arellano, A. (2008) 'False memory for associated word lists in individuals and collaborating groups', *Memory & cognition*, 36(3), pp. 598-603.
- Malloy, L.C., Wright, D.B. and Skagerberg, E.M. (2012) '5 Eyewitness Evidence', *Forensic Psychology: Crime, Justice, Law, Interventions*, pp. 99-114.
- Malpass, R.S. (1996) 'Enhancing eyewitness memory', *Psychological issues in eyewitness identification*, pp.177-204.
- Marshall, M.N. (1996) 'Sampling for qualitative research', *Family practice*, 13(6), pp. 522-526.
- Marshall, J.A., Brown, G. and Radford, A.N. (2017) 'Individual Confidence-Weighting and Group Decision-Making', *Trends in ecology & evolution*, 32(9), pp. 636-645.

- Martín, A.M. and Kaplan, M.F. (2006) 'Psychological Perspectives on Spanish and Russian Juries', in Kaplan, M.F. and Martín, A.M. (eds.) *Understanding World Jury Systems Through Social Psychological Research*. New York and Hove: Psychology Press, pp. 71-87.
- Maynard, P., Allwell, K., Roux, C., Dawson, M. and Royds, D. (2001) 'A protocol for the forensic analysis of condom and personal lubricants found in sexual assault cases', *Forensic science international*, 124(2-3), pp. 140-156.
- McBurney, D. and White, T.L. (2009) *Research Methods*. Belmont, Calif.: Wadsworth, Cengage Learning.
- McCartney, C. (2012) *Forensic Identification and Criminal Justice*. New York: Routledge.
- McCartney, C. and Amankwaa, A. (2017) 'DNA databases: it's still far from clear how effective they are in fighting crime', *The Conversation*.
- McDonald, S., Daniels, K. and Harris, C. (2004) 'Cognitive Mapping in Organizational Research', in Cassell, C. and Symon, G. (eds.) *Essential guide to qualitative methods in organizational research*. London; Thousands Oaks: SAGE Publications, pp. 73-86.
- McLellan, E., MacQueen, K.M. and Neidig, J.L. (2003) 'Beyond the Qualitative Interview: Data Preparation and Transcription', *Field Methods*, 15(1), pp. 63-84.
- McQuiston-Surrett, D. and Saks, M.J. (2007) 'Communicating opinion evidence in the forensic identification sciences: Accuracy and impact', *Hastings LJ*, 59, pp. 1159-1553.
- Meindl, J.R. (1995) 'The romance of leadership as a follower-centric theory: A social constructionist approach', *The leadership quarterly*, 6(3), pp. 329-341.
- Mercier, H. and Landemore, H. (2012) 'Reasoning is for arguing: Understanding the successes and failures of deliberation', *Political Psychology*, 33(2), pp. 243-258.
- Mercier, H. and Sperber, D. (2011) 'Why do humans reason? Arguments for an argumentative theory', *Behavioral and Brain Sciences*, 34(2), pp. 57-74.
- Mischel, W. (2007) 'Toward a cognitive social learning reconceptualization of personality', in Shoda, Y., Cervone, D. and Downey, G. (eds.) *Persons in Context: Building a Science of the Individual*. New York, London: Guilford, pp. 278-326.
- Moretti, F., van Vliet, L., Bensing, J., Deledda, G., Mazzi, M., Rimondini, M., Zimmermann, C. and Fletcher, I. (2011) 'A standardized approach to qualitative content analysis of focus group discussions from different countries', *Patient education and counseling*, 82(3), pp. 420-428.

- Morgan, D. (2013) *Integrating qualitative and quantitative methods: A pragmatic approach*. London: Sage publications.
- Morse, J.M. (2003) 'Principles of mixed methods and multimethod research design', in Tashakkori, A. and Teddlie, C. (eds.), *Handbook of Mixed Methods in Social and Behavioural Research*. Thousand Oaks: Sage, pp. 189-208.
- Moschella, G. (2009) *Magistratura e legittimazione democratica*. Milano: Giuffrè.
- Myers, M.A. (1979) 'Rule Departures and Making Law: Juries and Their Verdicts', *Law & Society Review*, 13(3), pp. 781-797.
- Myers, R.D., Reinstein, R.S. and Griller, G.M. (1999) 'Complex scientific evidence and the jury.(Genes and Justice)', *Judicature*, 83(3), pp. 150-156.
- Nance, D.A. and Morris, S.B. (2005) 'Juror understanding of DNA evidence: An empirical assessment of presentation formats for trace evidence with a relatively small random-match probability', *The Journal of Legal Studies*, 34(2), pp. 395-444.
- Newman, J. (1993) *Beyond "Reasonable Doubt"*. N.Y.: U.S. Court of Appeals.
- Nickerson, R.S. (1998) 'Confirmation Bias: A Ubiquitous Phenomenon in Many Guises', *Review of General Psychology*, 2(2), pp. 175-220.
- Noaks, L. and Wincup, E. (2004) *Criminological Research: Understanding Qualitative Methods*. United Kingdom: Sage Publications Ltd.
- Núñez, N., McCrea, S.M. and Culhane, S.E. (2011) 'Jury decision making research: Are researchers focusing on the mouse and not the elephant in the room?', *Behavioral sciences & the law*, 29(3), pp. 439-451.
- Oliver, R.L. (2014) *Satisfaction: A behavioral perspective on the consumer*. London: Routledge.
- Over, D. (2004) 'Rationality and the normative/descriptive distinction', *Blackwell handbook of judgment and decision making*, , pp. 3-18.
- Padoa Schioppa, A. (1994) *La giuria penale in Francia dai "Philosophes" alla Costituente*. Milano: LED.
- Parrott, C.T., Neal, T., Wilson, J.K. and Brodsky, S.L. (2015) 'Differences in expert witness knowledge: Do mock jurors notice and does it matter?', *Journal of the American Academy of Psychiatry and the Law*, 43(1), pp.69-81.
- Patton, M.Q. (2002) *Qualitative research and evaluation methods*. Thousand Oaks, Calif.; London: Sage.

- Pennington, N. and Hastie, R. (1981) 'Juror decision-making models: The generalization gap', *Psychological bulletin*, 89(2), pp. 246-287.
- Pennington, N. and Hastie, R. (1986) 'Evidence Evaluation in Complex Decision Making', *Journal of personality and social psychology*, 51(2), pp. 242-258.
- Pennington, N. and Hastie, R. (1988) 'Explanation-Based Decision Making: Effects of Memory Structure on Judgment', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3), pp. 521-533.
- Pennington, N. and Hastie, R. (1991) 'A cognitive theory of juror decision making: the story model', *Cardozo Law Review*, 13(2-3), pp. 519-557.
- Pennington, N., and Hastie, R. (1992). 'Explaining the evidence: Tests of the story model for juror decision making', *Journal of Personality and Social Psychology*, 62, 189-206.
- Pennington, N. and Hastie, R. (1993) 'Reasoning in explanation-based decision making', *Cognition*, 49(1), pp. 123-163.
- Pennington, N., and Hastie, R. (1994) 'The story model for juror decision making', in R. Hastie (Ed.), *Inside the juror*. New York: Cambridge University Press, pp. 192-224.
- Peoples, C.D., Sigillo, A.E., Green, M. and Miller, M.K. (2012) 'Friendship and conformity in group opinions: Juror verdict change in mock juries', *Sociological Spectrum*, 32(2), pp. 178-193.
- Pigott, M.A. and Foley, L.A. (1995) 'Social Influence in Jury Decision Making', *Trial Dipl.J.*, 18, pp. 101-108.
- Pizzorusso, A. (1990) *L'organizzazione della giustizia in Italia*. Roma: Einaudi.
- Ploscowe, M. (1935) 'The Development of Present-Day Criminal Procedures in Europe and America', *Harvard law review*, 48(3), pp. 433-473.
- Polkinghorne, D.E. (2005) 'Language and Meaning: Data Collection in Qualitative Research', *Journal of Counseling Psychology*, 52(2), pp. 137-145.
- Pollock, F. and Maitland, F.W. (1895) *The history of English law before the time of Edward 1*. Cambridge: Cambridge University Press.
- Popov, V. and Reder, L. (2017) *Repetition improves memory by strengthening existing traces: Evidence from paired-associate learning under midazolam*. Available at: https://www.researchgate.net/profile/Vencislav-Popov/publication/317039436_Repetition_improves_memory_by_strengthening_existing_traces_Evidence_from_paired-associate_learning_under_midazolam/links/5921ed7d458515e3d4076a01/Repeti

[tion-improves-memory-by-strengthening-existing-traces-Evidence-from-paired-associate-learning-under-midazolam.pdf](#) (Accessed: 27/03/2018).

- Poynter, R. (2001) 'A guide to best practice in online quantitative research', *The challenge of the Internet*, pp. 3-19.
- Pritchard, M.E. and Keenan, J.M. (2002) 'Does jury deliberation really improve jurors' memories?', *Applied Cognitive Psychology*, 16(5), pp. 589-601.
- Punch, K.F. (1998) *Introduction to social research: quantitative and qualitative approaches*. London: Sage.
- Punch, K.F. (2000) *Developing effective research proposals*. London: Sage.
- Pyett, P.M. (2003) 'Validation of qualitative research in the "real world"', *Qualitative health research*, 13(8), pp. 1170-1179.
- Reifman, A., Gusick, S.M. and Ellsworth, P.C. (1992) 'Real jurors' understanding of the law in real cases', *Law and human behavior*, 16(5), pp. 539-554.
- Renner, M. and Taylor-Powell, E. (2003) 'Analyzing qualitative data', *Programme Development & Evaluation, University of Wisconsin-Extension Cooperative Extension*, pp. 1-10.
- Rohner, R.P. (1977) 'Advantages of the comparative method of anthropology', *Behavior Science Research*, 12(2), pp. 117-144.
- Rose, V.G. and Ogloff, J.R.P. (2001) 'Evaluating the comprehensibility of jury instructions: a method and an example', *Law and Human Behavior*, 25(4), pp. 409-431.
- Ruva, C., McEvoy, C. and Bryant, J.B. (2007) 'Effects of pre-trial publicity and jury deliberation on juror bias and source memory errors', *Applied Cognitive Psychology*, 21(1), pp. 45-67.
- Ryan, G.W. and Bernard, H.R. (2000) 'Data management and analysis methods', in Densin, N. and Lincoln, Y. (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications, pp. 769-802.
- Saks, M.J. and Faigman, D.L. (2008) 'Failed forensics: How forensic science lost its way and how it might yet find it', *Annual Review of Law and Social Science*, 4, pp. 149-171.
- Salerno, J.M. and Diamond, S.S. (2010) 'The promise of a cognitive perspective on jury deliberation', *Psychonomic bulletin & review*, 17(2), pp. 174-179.
- Santalucia, B. (2010) *Altri studi di diritto penale romano*. Wolters Kluwer Italia.

- Saunders, H.D. (2005) *Quantifying reasonable doubt: A proposed solution to an equal protection problem*. Bepress Legal Series.
- Scandura, T.A. and Williams, E.A. (2000) 'Research Methodology in Management: Current Practices, Trends, and Implications for Future Research', *The Academy of Management Journal*, 43(6), pp. 1248-1264.
- Schklar, J. and Diamond, S.S. (1999) 'Juror reactions to DNA evidence: Errors and expectancies', *Law and Human Behavior*, 23(2), pp.159-184.
- Schmechel, R.S., O'Toole, T.P., Easterly, C. and Loftus, E.F. (2006) 'Beyond The Ken? Testing Jurors' Understanding Of Eyewitness Reliability Evidence', *Jurimetrics*, 46(2), pp. 177-214.
- Schweitzer, N.J. and Saks, M.J. (2007) 'The CSI Effect: popular fiction about forensic science affects the public's expectations about real forensic science', *Jurimetrics Journal of Law, Science and Technology*, 47(3), pp. 357-364.
- Severance, L.J. and Loftus, E.F. (1982) 'Improving the Ability of Jurors to Comprehend and Apply Criminal Jury Instructions', *Law & Society Review*, 17(1), pp. 153-197.
- Shafir, E., Simonson, I. and Tversky, A. (1993) 'Reason-based choice', *Cognition*, 49(1), pp. 11-36.
- Shelton, S.T. (2006) 'Jury Decision Making: Using Group Theory to Improve Deliberation', *Politics & Policy*, 34(4), pp. 706-725.
- Sherif, M. (1936) *The psychology of social norms*. New York: Octagon Books.
- Shuman, D.W. and Champagne, A. (1997) 'Removing the people from the legal process: The rhetoric and research on judicial selection and juries.', *Psychology, Public Policy, and Law*, 3(2-3), pp. 242-258.
- Silverman, D. (1993) *Interpreting qualitative data: Methods for analysing talk, text and interaction*. London: SAGE.
- Simon, D. (2004) 'A third view of the black box: Cognitive coherence in legal decision making', *The University of Chicago Law Review*, pp. 511-586.
- Simon, D., Pham, L.B., Le, Q.A. and Holyoak, K.J. (2001) 'The emergence of coherence over the course of decision making', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(5), pp. 1250-1260.
- Simon, D., Snow, C.J. and Read, S.J. (2004) 'The redux of cognitive consistency theories: evidence judgments by constraint satisfaction.', *Journal of personality and social psychology*, 86(6), pp. 814-837.

- Simon, R.J. and Mahan, L. (1971) 'Quantifying burdens of proof: A view from the bench, the jury, and the classroom', *Law and Society Review*, 5(3), pp. 319-330.
- Simonson, I. (1989) 'Choice based on reasons: the case of attraction and compromise effects', *Journal of Consumer Research*, 16(2), pp. 158-174.
- Smith, V.L. (1991) 'Prototypes in the courtroom: Lay representations of legal concepts', *Journal of personality and social psychology*, 61(6), pp. 857-872.
- Smith, V.L. (1993) 'When prior knowledge and law collide: Helping jurors use the law', *Law and human behavior*, 17(5), pp. 507-536.
- Smith, V.L. and Studebaker, C.A. (1996) 'What do you expect?: The influence of people's prior knowledge of crime categories on fact-finding', *Law and human behavior*, 20(5), pp. 517-532.
- Smith, E.R., Mackie, D.M. and Claypool, H.M. (2014) *Social psychology*. Philadelphia, Pa.; Hove: Psychology Press.
- Smith, L.L., Bull, R. and Holliday, R. (2011) 'Understanding juror perceptions of forensic evidence: investigating the impact of case context on perceptions of forensic evidence strength', *Journal of forensic sciences*, 56(2), pp. 409-414.
- Spencer, S.E., Kim, S.Y., Kim, S.B. and Schug, K.A. (2011) 'Matrix-assisted laser desorption/ionization–time of flight-mass spectrometry profiling of trace constituents of condom lubricants in the presence of biological fluids', *Forensic science international*, 207(1-3), pp. 19-26.
- Staveley, E.S. (1954) 'Provocatio during the Fifth and Fourth Centuries B. C.', *Historia*, 3, pp. 412-428.
- Staw, B.M. (1975) 'Attribution of the “causes” of performance: A general alternative interpretation of cross-sectional research on organizations', *Organizational behavior and human performance*, 13(3), pp. 414-432.
- Stebay, N., Hosch, H.M., Culhane, S.E. and McWethy, A. (2006) 'The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis', *Law and human behavior*, 30(4), pp. 469-492.
- Steckler, A., McLeroy, K.R., Goodman, R.M., Bird, S.T. and McCormick, L. (1992) 'Toward integrating qualitative and quantitative methods: an introduction', *Health Education & Behavior*, 19(1), pp. 1-8.
- Stewart, D.W. and Shamdasani, P.N. (2014) *Focus groups: Theory and practice*. London: Sage publications.
- Stolfi, E. (2006) *Introduzione allo studio dei diritti greci*. Torino: Giappichelli.

- Strauss, A. and Corbin, J. (1990) *Basics of Qualitative Research*. Newbury Park, CA: Sage.
- Suhonen, N., 2007. 'Normative and descriptive theories of decision making under risk: A short review'. *Finland: Economics and Business Administration*, University of Joensuu, Finland.
- Summers, R.S. (1999) 'Formal Legal Truth and Substantive Truth in Judicial Fact-Finding--Their Justified Divergence in some Particular Cases', *Law and Philosophy*, 18(5), pp. 497-511.
- Tarr, G.A. (2013) *Judicial process and judicial policymaking*. Cengage Learning.
- Taylor, G.W. and Ussher, J.M. (2001) 'Making Sense of S&M: A Discourse Analytic Account', *Sexualities*, 4(3), pp. 293-314.
- Temkin, J. (2000) 'Prosecuting and defending rape: Perspectives from the bar', *Journal of Law and Society*, 27(2), pp. 219-248.
- Tetlock, P.E. (1985) 'Accountability: The neglected social context of judgment and choice', *Research in organizational behavior*, 7(1), pp. 297-332.
- Thagard, P. and Millgram, E. (1995) 'Inference to the Best Plan: A Coherence Theory of Decision', *Goal-driven learning*, pp. 439-454.
- Tinsley, Y. (2001) 'Juror decision-making: a look inside the jury room', *Selected proceedings of the British Society of Criminology Conference*, 4.
- Loftu and Tonini, P. (2011) *Diritto processuale penale. Manuale breve. Tutto il programma d'esame con domande e risposte commentate*. Milano: Giuffrè Editore.
- Tversky, A. and Kahneman, D. (1974) 'Judgment under Uncertainty: Heuristics and Biases', *Science (New York, N.Y.)*, 185(4157), pp. 1124-1131.
- Vico, P. (1904) 'Giurati', in *Il Digesto italiano*. Torino: Unione tipografica editrice torinese, p. 563.
- Vidmar, N. (1994) 'Are Juries Competent to Decide Liability in Tort Cases Involving Scientific/Medical Issues--Some Data From Medical Malpractice', *Emory Law Journal*, 43(3), pp. 885-911.
- Vidmar, N. (2000) *World jury systems*. Oxford: Oxford University Press.
- Vidmar, N. (2008) 'Civil juries in ecological context: Methodological implications for research', in *Civil Juries and Civil Justice*, Springer, pp. 35-65.

- Vollrath, D.A., Sheppard, B.H., Hinsz, V.B. and Davis, J.H. (1989) 'Memory performance by decision-making groups and individuals', *Organizational behavior and human decision processes*, 43(3), pp. 289-300.
- Vroom, V.H. and Yetton, P.W. (1973) *Leadership and decision-making*. University of Pittsburgh Press.
- Wagenaar, W.A., van Koppen, P.J. and Crombag, H.F. (1993) *Anchored narratives: The psychology of criminal evidence*. St Martin's Press.
- Wasamba, P. (2015) *Contemporary oral literature fieldwork*. University of Nairobi Press.
- Whalen, T. and Churchill, G. (1971) 'Decisions under uncertainty', Robinson College of Business Georgia State University.
- Whitworth, B. and McQueen, R.J. (2003) 'Voting before discussing: Electronic voting as social interaction', *Group Facilitation*, (5), pp. 1-20.
- Wiener, R.L., Hurt, L.E., Thomas, S.L., Sadler, M.S., Bauer, C.A. and Sargent, T.M. (1998) 'The Role of Declarative and Procedural Knowledge in Capital Murder Sentencing¹', *Journal of Applied Social Psychology*, 28(2), pp. 124-144.
- Wiener, R.L., Krauss, D.A. and Lieberman, J.D. (2011) 'Mock jury research: where do we go from here? (Jury Simulation Research: Student Versus Community Samples)', *Behavioral sciences & the law*, 29(3), pp. 467-479.
- Willmott, D., Boduszek, D., Debowska, A. and Woodfield, R. (2018) 'Introduction and validation of the Juror Decision Scale (JDS): An empirical investigation of the Story Model', *Journal of Criminal Justice*, 57, pp. 26-34.
- Wilson, D.W. and Donnerstein, E. (1977) 'Guilty or not guilty? A look at the "simulated" jury paradigm', *Journal of Applied Social Psychology*, 7(2), pp. 175-190.
- Winter, R. and Greene, E. (2007) 'Juror Decision-Making', in Durso, F.T., Nickerson, R.S., Dumais, S.T., Lewandowsky, S. and Perfect, T.J. (eds.) *Handbook of Applied Cognition*. Chichester, UK: John Wiley & Sons Ltd, pp. 739-761.
- Wolff, C. (1994) 'Rapists and Condoms; Is Use a Cavalier Act or a Way to Avoid Disease and Arrest?', *The New York Times*.
- Wolman, B. (1956) 'Leadership and group dynamics', *The Journal of social psychology*, 43(1), pp. 11-25.
- Woodhams, J. and Labuschagne, G. (2012) 'South African serial rapists: The offenders, their victims, and their offenses', *Sexual Abuse*, 24(6), pp. 544-574.

- Wu, G., Zhang, J. and Gonzalez, R., 2004. 'Decision under risk'. *Blackwell handbook of judgment and decision making*, pp. 399-423.
- Zambuto, F. (2016) 'L'ombra della "sentencing disparity": casi giudiziari analoghi e conclusioni differenti', *Cammino Diritto - Rivista di informazione giuridica*, 2(ISSN 2532-9871), pp. 1-18.
- Zaragoza, M.S., Belli, R.F. and Payment, K.E. (2007) 'Misinformation effects and the suggestibility of eyewitness memory', in Garry, M. and Hayne, H. (eds.) *Do Justice and Let the Sky Fall: Elizabeth F. Loftus and Her Contributions to Science, Law and Academic Freedom*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 35-63.
- Zeisel, H. and Diamond, S.S. (1978) 'The effect of peremptory challenges on jury and verdict: An experiment in a federal district court', *Stanford law review*, 30(3), pp. 491-531.