‘Lock them up and throw away the key’: an evaluation of the structure of punitive attitudes

Nicolás Trajtenberg, Pablo Ezquerra & Matthew Williams

To cite this article: Nicolás Trajtenberg, Pablo Ezquerra & Matthew Williams (07 Apr 2024): ‘Lock them up and throw away the key’: an evaluation of the structure of punitive attitudes, Psychiatry, Psychology and Law, DOI: 10.1080/13218719.2023.2296476

To link to this article: https://doi.org/10.1080/13218719.2023.2296476

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 07 Apr 2024.

Submit your article to this journal

View related articles

View Crossmark data
‘Lock them up and throw away the key’: an evaluation of the structure of punitive attitudes

Nicolás Trajtenberg, Pablo Ezquerre and Matthew Williams

1. Introduction

Punishment as a response to deviance and crime is a ubiquitous practice in all societies throughout history (Miethe & Lu, 2005). There has been a growing interest in citizens’ attitudes towards punishment (Adriaenssen & Aertsen, 2015; Aguilar-Jurado, 2018), particularly in the context of a ‘punitive turn’ observed in western societies in the last decades (Downes, 2011; Enns, 2014; Sozzo, 2018; Unnever & Cullen, 2010). Punitive attitudes might involve the support of harsher penal punishments and can have real consequences on the functioning of the criminal justice system and crime prevention policies (Enns, 2014; Jennings et al., 2017). Instead of following expert advice and empirical evidence, populist politicians might be more guided by the opinions and emotional reactions of the public (Bottoms, 1995; Garland, 2021). Some of these policies that involve greater use of imprisonment and longer sentences not only do not work but might involve several negative externalities in terms of recidivism, human rights and economic costs (Cullen et al., 2011; Liu et al., 2018; Loeffler & Nagin, 2022). Additionally, even assuming that citizens’ opinions should be considered as a limit to expert knowledge (Robinson, 2012), it is not clear how politicians actually know these punitive preferences. Usually, they are more imagined or based on weak surveys than based on democratic consultation or robust empirical evidence (Gargarella, 2019; Garland, 2021). In fact, some research shows that when public
penal attitudes are more carefully considered, they are less punitive and closer to those of policymakers and criminal justice professionals (Cullen et al., 1988; J. V. Roberts & Stalans, 2019).

Despite the relevance of the topic and the fact that there have been more than three decades of empirical research, there is still little agreement in the literature regarding the definition of punitiveness, its main components and how we should measure it (Adriaenssen & Aertsen, 2015; Maguire & Johnson, 2015; Matthews, 2005). Empirical research has focused on explaining punitive attitudes and identifying their main economic, social, political and cultural determinants (Kleck & Jackson, 2017; Lehmann & Pickett, 2017; Unnever & Cullen, 2010) rather than assessing their validity, content and measurement structure (Aizpurúa, 2015; Maguire & Johnson, 2015). As a result, most studies involve the multivariate analysis of the determinants of punitiveness using either single item measures – such as support for the death penalty (e.g. Lehmann & Pickett, 2017) – or global indexes or scores of punitive questions (e.g. Chiricos et al., 2004; Maruna & King, 2009; Spiranovic et al., 2012). These studies often do not clearly state the dimensionality of punitiveness, let alone testing whether there is a common underlying punitive construct. Some relevant research has started to evaluate the structure of punitive attitudes to understand whether this is a single construct or rather a more heterogeneous phenomenon (Aizpurúa, 2015; Armborst, 2017; Maguire & Johnson, 2015; Mascini & Houtman, 2006; Ortet-Fabregat & Pérez, 1992; Silver & Silver, 2017). Despite this being a relevant initial step, analysis is still based on simple exploratory and confirmatory techniques and runs the risk of providing a biased picture of the nature of punitive attitudes.

The rest of the paper is ordered as follows. We first discuss the challenges of conceptualisation and measurement of punitive attitudes in criminology. Then we describe previous empirical studies that have evaluated the construct’s structure. The third section describes the data, measures and type of statistical analysis used in this paper. The fourth section presents the main findings, and the final section discusses our results, the main limitations and future lines of research.

2. Conceptualisation and measurement of punitive attitudes

2.1. Disagreement in the conceptualisation and measurement of punitiveness

The concept of punitive attitudes remains ‘under-theorised’ and vague (Adriaenssen & Aertsen, 2015; Carvalho et al., 2020; Healy & McGrath, 2019; Pfeffer, 2023), and there are still strong disagreements in the field regarding: what the different levels of analysis of punitiveness are; how it should be conceptualised; how multifaceted it is and which kinds of dimensions or components should be considered (e.g. what is the role of sentiments, beliefs, goals, or what is the target of punitiveness); or how it should be measured and how global or specific its operationalisation should be (Burton et al., 2020; Hamilton, 2014; Kornhauser, 2015; Kury et al., 2009; Matthews, 2014; Pickett, 2019). On conceptualisation and measurement there are several key disagreements. One first conceptual issue concerns what kind of specific mental states are involved in punitiveness. Surprisingly, many studies in the field of punitiveness do not even include an explicit definition of the key concept (e.g. Aizpurúa, 2015; Pickett & Baker, 2014) or even operational definitions (probability of endorsing punishment as a response to crime; e.g. Sargent, 2004). Some authors use vague and general terms such as ‘support’ (e.g. Baker et al., 2015; Unnever & Cullen, 2009), or at best some studies focus on support for policies without clearly defining the aspects of these policies that render them punitive beyond their opposition to rehabilitation (Pickett & Baker, 2014; Ramirez, 2013). Some authors almost tautologically define punitive attitudes simply as ‘attitudes toward
The Structure of Punitive Attitudes

sanctioning or punishment’ (e.g. Mackey & Courtright, 2000), while some refer to a general attitude that considers that punishment in its actual form is not severe or intense enough (Barrett et al., 2023). Other authors use more specific mental states such as ‘sentiments’ (e.g. Ramirez, 2013; Stewart et al., 2018), ‘desires’ (e.g. Spiranovic et al., 2012) or ‘preferences’ (e.g. Baker et al., 2015). Finally, there are cases of catch-all definitions where punitive attitudes are ‘attitudes’, ‘emotions’, ‘sentiments’ and ‘style of reaction to crime’ (Armborst, 2017). Despite the significant differences between all these terms, in most cases, little argument is provided for selection. Most of the studies use the term ‘attitudes’, but only Maguire and Johnson (2015) explicitly discuss the concept in the context of cognitive and social psychology. According to this literature, punitive attitudes can be considered as one specific type of evaluative attitude (Ajzen & Fishbein, 2005; Albarracin & Shavitt, 2018; Riemer et al., 2014) – that is, an individual tendency to have a positive or approval reaction to a specific psychological object: criminal justice policies or institutions that promote the increase of the costs of crime for offenders through the application of punishments (Aguilar-Jurado, 2018; Brooks, 2021). While achieving a universally agreed-upon definition of punitive attitudes may not be feasible or even desirable, there is clear utility in discussing how to enhance the precision and coherence of the concept (Gerring, 2011).

Lack of clarity on the concept of punitive-ness involves not only the nature of mental states but also the content of those mental states. Disagreements involve not only the actual meaning of the term punitive, and whether it involves imposing higher penal costs, how excessive they need to be, or even the role of punishment reasons such as rehabilitation or retribution (Matthews, 2005, 2014), but also what is the adequate operationalisation of these evaluative attitudes (Adriaenssen & Aertsen, 2015; Aizpurua, 2015). When it comes to measuring these punitive attitudes, very different indicators have been used. First, some studies include measures of citizens’ support for specific stiffer penal sanctions or sentencing policies (Gerber & Jackson, 2016; Hogan et al., 2005; Pickett & Baker, 2014), particularly the preference for death penalty (Cullen et al., 1985; Silver & Silver, 2017), long-term or life prison sentences (Armborst, 2017; Mackey & Courtright, 2000) and mandatory minimum sentences (Silver & Silver, 2017). Second, studies include measures of a general inclination for a less benevolent or lenient criminal justice system that should punish more severely (Grasmick & McGill, 1994; Maguire & Johnson, 2015; Maruna & King, 2009), or where punishments should have a more central role (Miller, 2014). Third, some scales include items that focus on prison conditions and the pains suffered by inmates, asking respondents whether prisons should have harsher conditions, lack of access to privileges or even degrading circumstances (Armborst, 2017; Hanslmaier & Baier, 2016; Hogan et al., 2005; Mackey & Courtright, 2000).

A fourth strategy involves including measures of the goals or justifications of penal sanctions. There are two main justifications for punishment: on the one hand, deontological or retributive reasons, where punishment is an end in itself and involves a ‘proportional’ response to the offender to restore the moral balance broken when an offence has taken place (Von Hirsch, 1998); and on the other hand, teleological or consequential arguments where punishment is a means to an end (reduction of crime or recidivism; Duff & Garland, 1994). Consequential arguments involve specific justifications: dissuading actual offenders from future recidivism (individual deterrence) or the general population from getting involved in criminal behaviour in the first place (general deterrence); incapacitation of offenders by limiting their freedom through imprisonment or, at its most extreme, by capital punishment; rehabilitation of offenders through psychological
treatment and training to help them to return to the community; conflict resolution between offender, victim and related parties to restore or restitute harms produced by crime; and finally reducing crime through alternatives to penal punishment that involve early prevention programmes in families or neighbourhoods (Canton, 2017; Duff & Garland, 1994; Oswald et al., 2002). Some studies include punishment goals in their scales and usually define respondents’ punitiveness as a rejection of rehabilitation and psychological treatment (and sometimes restoration and social prevention) and/or the support of retribution, incapacitation (and sometimes deterrence; Cullen et al., 1988; Mascini & Houtman, 2006; L. D. Roberts & Indermaur, 2007). However, some authors consider punishment goals a separate construct that needs to be related to a punitive attitudes construct (Mascini & Houtman, 2006; Oswald et al., 2002; Payne et al., 2004; Pickett & Baker, 2014).

An additional indicator used is attitudes regarding police and policing. Punitiveness is captured by evaluating respondents’ support for increasing the amount of police on the streets (Costelloe et al., 2009; Ortet-Fabregat & Pérez, 1992), approving police being tough (Armbrorst, 2017), endorsing police abuse, breaking the law, or even approving licence to kill criminals as a better way to control crime (Maguire & Johnson, 2015; see also Mascini & Houtman, 2006). A sixth indicator includes punishment outside of the criminal justice – that is, items that tap into respondents’ approval of vigilantism in the community (Borraz et al., 2012; Tyler & Boeckmann, 1997). Finally, some studies measure punitiveness by applying some of the aforementioned indicators to a particularly vulnerable group, such as juvenile offenders, and thus assessing respondents’ willingness to make sentences more severe, send them to prison, prosecute them as adults or even to lower the age of criminal responsibility (Chiricos et al., 2004; Costelloe et al., 2009; Piquero et al., 2010).

2.2. Multidimensionality or unidimensionality of punitiveness

Given the array of indicators discussed in the previous section, a relevant issue is whether punitive attitudes constitute a single construct or, instead, a multidimensional one. In this scenario, it is possible that the objectives of punishment are a component of an overall punitive attitude, which incorporates other dimensions, such as the type and severity of the punishment.

According to Maguire and Johnson (2015), there are two approaches to treating punitive attitudes as a unidimensional construct: (a) scales composed of items where the highest scores involve strong support for more punitive policies, and the lowest scores involve strong support for progressive policies such as rehabilitation or restoration; (b) scales that include items that focus only on the evaluation of punitive policies/sentencing ranging from high support to weak support, excluding items on progressive policies (Maguire & Johnson, 2015; see also Mascini & Houtman, 2006). Both solutions are problematic and provide a biased estimation of the structure of punitive attitudes: the first strategy falls into a ‘double-barrelled-question’ problem (Oppenheim, 2000) since it assumes that punitive and progressive policies are polar and inconsistent opposites and forces respondents to choose between both options; the second strategy captures a better representation of approval of punitive policies, but without including questions regarding progressive policies it can lead to the misunderstanding that individuals who strongly support punitive policies will necessarily reject policies focused on rehabilitation or social prevention (Cullen et al., 2000;

\footnote{Some scales include items that measure respondents’ support of weakening courts and judges’ powers to obstruct the work of police (Ortet-Fabregat and Pérez, 1992) or to defend the rights of offenders (Viney et al., 1982).}
Hutton, 2005; Maguire & Johnson, 2015; Mascini & Houtman, 2006). For many authors, instead, punitiveness is a complex, ambivalent and multifaceted phenomenon where punitive and more progressive views are both part of our neuropsychological human nature (Maguire & Johnson, 2015; see also Bloom, 2017), but they cannot be reconciled or reduced to extremes of a single construct (Cullen et al., 2000; Duffee, 1980; Hutton, 2005). However, there are disagreements about what kind of relationship we should expect between these different dimensions. While some authors argue for an inverse relationship or compensatory hypothesis where higher scores on punitive attitudes imply lower scores on progressive ones and vice versa (Ortet-Fabregat & Pérez, 1992; Pickett et al., 2013), others argue for a positive relation where individuals might have simultaneously punitive and progressive attitudes (Cullen et al., 2007; Falco & Turner, 2014; Maguire & Johnson, 2015). Finally, others argue for progressive attitudes, particularly regarding rehabilitation goals, to have an inconsistent relation with punitive attitudes (Mascini & Houtman, 2006).

The result of this wide variety of different indicators and lack of common criteria is a very inconsistent operationalisation of the concept of punitive attitudes (Adriaenssen & Aertsen, 2015; Aizpurúa, 2015). This dispersion and heterogeneity of measures demand a more appropriate psychometric measurement of punitive attitudes’ unidimensional or multidimensional structure.

3. Previous studies
The conceptual discussion regarding the content and dimensionality of punitive attitudes requires empirical evaluation. However, psychometric evaluation is scarce and, at best, is conducted using exploratory analysis and, only more recently, some confirmatory techniques.

Two studies that used only exploratory techniques found empirical support for the multidimensionality of punitive attitudes. An early study that focused on analysing attitudes toward crime among professionals of the criminal justice system in Spain using a 25-item scale (Attitudes toward the Prevention of Crime Scale) found support in terms of reliability and stability for two separate dimensions, one regarding coercive prevention mainly based on deterrence and one regarding social prevention based on intervention through social agencies and the community (Ortet-Fabregat & Pérez, 1992). Likewise, Mascini and Houtman’s (2006) study on the integration between rehabilitation and repression values used a representative Dutch sample to distinguish three constructs regarding attitudes of support for repression (6 items), support for rehabilitation (12 items) and support for decriminalisation (6 items). Yet, validity and reliability tests were performed only for each of the scales without testing the existence of a connection between items and/or a potential latent construct.

However, exploratory analysis research has also found support for the unidimensionality of attitudes toward punishment. Ortet-Fabregat and Pérez’s (1992) study also included the exploratory analysis of a 22-items scale (Attitudes towards the Treatment of Crime Scale), which shows support for a single factor (called assistance vs. punishment) that combines both items associated with severe punishment of offenders (e.g. physical punishment and forced labour) with more rehabilitative, educative and alternative to prison measures. Another study conducted by Piquero et al. (2010) on public opinion optimism about juvenile rehabilitation and its
socio-demographic correlates also found evidence for a single factor in a 3-item scale with items focused on support for rehabilitation of juvenile offenders in an American sample.

More recently, some studies have assumed a more deductive approach to measurement, including confirmatory techniques but with contradictory results. A study that sought the elaboration and validation of punitive scales in a Spanish sample of male and female adults used a combination of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) to show the reliability and validity of a single factor of punitive attitudes both for a 7-item attitude toward crime scale and for a 10-item attitude toward juvenile delinquency scale, which included support for harsher sentences, control of the streets, prison conditions, prisons’ perverse effects on recidivism, support for early social prevention, and so on (Aizpurúa, 2015). A study using a German sample focused on understanding the role of fear of crime as predictor of punitiveness applied CFA and also found empirical support for a single factor of punitiveness that comprised severity of punishment, including death penalty, punishment for juvenile offenders and sexual offenders, age of penal responsibility, and so on (Armborst, 2017). However, according to Armborst (2017), the robustness of the construct is problematic given that the estimation process involved disregarding several relevant survey items. A multisite study in seven Caribbean countries by Maguire and Johnson (2015), which focused explicitly on the evaluation of the multidimensionality of punitive attitudes using EFA and CFA in a battery of 11 punitive and progressive items, also found little support for unidimensionality, with three different factors being found. While items related to support for rehabilitation and social prevention loaded into a single progressive factor, punitive items responded to two different constructs: punitiveness and support for extra-judicial measures. Similarly, a study by Ramirez (2015) examining change over time in support for punitive anti-crime policies in a sample of black Americans used CFA to demonstrate that punitive attitudes involved two different constructs exhibiting a significant negative correlation between the two: support for crime prevention measures, which involved items such as tougher criminal sentencing or stricter parole conditions; and support for crime programmes with questions associated to a preference for educational or community programmes. Finally, a study that used the Moral Foundation Approach (Graham et al., 2009) to evaluate the role between group-oriented moral and ideological values and punitive attitudes assessed the psychometric structure of punitive attitudes in two North American samples (Silver & Silver, 2017). The first sample evaluated Chiricos et al.’s (2004) 8-item punitive scale on a sample of undergraduate students, finding little support for unidimensionality since EFA indicated three factors associated with punitive attitudes (support for tougher sanctions for adults and juveniles, support for the death penalty and support for tougher treatment during incarceration), and CFA showed a poor fit. The second sample evaluated a 6-item scale adapted from Pickett and Baker (2014) on an online sample of adults finding mixed support: while EFA revealed that items representing the reversed and non-reversed sets of items, loaded on two different factors, CFA showed that the unidimensional scale was an excellent fit for the data (Table 1).

These recent studies reveal a growing interest in the evaluation of the structure and dimensionality of items measuring punitive attitudes. However, the results obtained are limited given the nature of the psychometric analysis conducted. Particularly, the conclusions regarding the multidimensionality of punitive attitudes run the risk of being a methodological artefact due to the type of statistical analysis conducted. The use of bifactor techniques is useful for examination of multifaceted constructs since it allows a more adequate testing of subdimensions within an overall dimension as an alternative to second-order models (Bornovalova...
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Sample characteristics</th>
<th>Method</th>
<th>Dimensions</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ortet-Fabregat and Pérez (1992)</td>
<td>Study 1 (n = 382) Spain Study 2 (n = 357) Spain</td>
<td>Study 1: Second-year university students at the Autonomous University of Barcelona. 72% females. $M_{age} = 20$ Study 2: Different groups of professionals in the criminal justice system from Catalonia, Spain, together with the student sample of Study 1. 46% females. $M_{age} = 26.4$</td>
<td>Alpha + EFA</td>
<td>Prevention: coercive + social intervention Treatment: assistance vs. punishment</td>
<td>Prevention: multidimensional Treatment: unidimensional</td>
</tr>
<tr>
<td>Mascini and Houtman (2006)</td>
<td>n = 1982 Netherlands</td>
<td>Representative sample of Dutch Population. 48.4% females. $M_{age}$ not reported.</td>
<td>Alpha + EFA</td>
<td>Support for prevention Support for rehabilitation Support for decriminalisation</td>
<td>Multidimensional</td>
</tr>
<tr>
<td>Piquero et al. (2010)</td>
<td>n = 1502 USA</td>
<td>Representative sample of Pennsylvania households. 59.7% females. $M_{age} = 50.2$</td>
<td>Corr. + Alpha + EFA</td>
<td>Global (juvenile rehabilitation)</td>
<td>Unidimensional</td>
</tr>
<tr>
<td>Maguire and Johnson (2015)</td>
<td>n = 11,155 Antigua and Barbuda, Barbados, Guyana, Jamaica, Saint Lucia, Suriname, and Trinidad and Tobago</td>
<td>Stratified representative sample of the seven countries. Mean age or gender not reported.</td>
<td>Corr. + EFA + CFA</td>
<td>Punitive Extra-judicial Progressive</td>
<td>Multidimensional</td>
</tr>
<tr>
<td>Aizpurúa (2015)</td>
<td>n = 1000 Spain</td>
<td>Representative sample of the Spanish population.</td>
<td>CFA</td>
<td>Punitive youth Punitive adults (differentiated for each population)</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Reference</th>
<th>Sample</th>
<th>Sample characteristics</th>
<th>Method</th>
<th>Dimensions</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramirez (2015)</td>
<td>n = 515 USA</td>
<td>Representative sample on the black population in the United States. Mean age or gender not reported.</td>
<td>Alpha + CFA</td>
<td>Punitive Preventive</td>
<td>Multidimensional</td>
</tr>
<tr>
<td>Armborst (2017)</td>
<td>n = 1272 Germany</td>
<td>Representative sample of the adult population in Germany. No mean age or gender reported.</td>
<td>CFA</td>
<td>Global (severity, death penalty, prison conditions, age of criminal responsibility, sex offenders)</td>
<td>Unidimensional</td>
</tr>
<tr>
<td>Silver and Silver (2017)</td>
<td>Study 1 (n = 1464) USA Study 2 (n = 1025). USA</td>
<td>Study 1: Undergraduate students at the Pennsylvania State University. 58% females. $M_{age} = 19.8$. Study 2: Amazon Mechanical Turk paid survey. 54% of females. $M_{age} = 37.0$.</td>
<td>Alpha + EFA + CFA</td>
<td>Study 1: harsher sanctions for adults and juveniles + death penalty + harsher treatment during incarceration Study 2: Global (adult offenders, juvenile offenders, and sex offenders).</td>
<td>Unidimensional (mixed evidence)</td>
</tr>
</tbody>
</table>

Note: EFA = exploratory factor analysis; CFA = confirmatory factor analysis; Alpha = Cronbach’s alpha; Corr. = Correlation analysis.
et al., 2020; Dombrowski et al., 2019). By separating the variance attributable to specific subdimensions more clearly from the variance attributable to the overall factor, it allows: a better understanding of the specific contribution of each subdimension and a more nuanced view of the core structure of the construct; and enhanced validity and reliability by ensuring that the general construct is measured independently of the specific subdimensions (Reise et al., 2010). Moreover, this technique includes fit indices to specifically evaluate the proportion of common variance that belongs to the general factor, providing an additional measure to evaluate the unidimensionality of the construct’s structure (Rodriguez et al., 2016).

4. Current study
In the present study, we focus on examining the structure of punitive attitudes using survey data from a South American nation and particularly focus on whether using different analytical approaches significantly changes results. Our goal is not to validate a new measure of punitiveness but rather to evaluate the structure of punitive attitudes using well-known and validated items of the literature. Our main hypothesis is that the type of statistical analysis will significantly affect the results obtained, namely whether the structure of punitive attitudes is multidimensional or unidimensional. Particularly, we believe that when more sophisticated analyses are applied, namely bifactor analysis, it is sound to describe punitive attitudes as a single or unidimensional construct, including not only items that tap into the severity or intensity of punishment but also punitive and progressive reasons. We will empirically evaluate the hypothesis using survey data from a representative sample, which includes 19 items tapping on different dimensions of punitive attitudes. We use an ample definition of punitive attitudes as an evaluative attitude that involves cognitive components associated with three specific dimensions that reflect most relevant aspects in the literature: not only severity or intensity of penal punishment (Barrett et al., 2023; Matthews, 2005), but also reasons for punishment, both punitive and progressive (Cullen et al., 1988). Additionally, our operationalisation also follows the literature by including not only attitudes towards severe sentencing policies and criminal justice systems but also a variety of goals or justifications to punish crimes, both deontological and consequentialist (Cullen et al., 1988, 2000; Mackey & Courtright, 2000), attitudes toward specific actors of the criminal justice system, such as courts and police (Maguire & Johnson, 2015; Ortet-Fabregat & Pérez, 1992), and specifically how these attitudes change when it comes to punishing juveniles (Piquero et al., 2010).

5. Method
5.1. Data
The data set used in this study comes from a cross-sectional telephone survey of 895 citizens in Uruguay that was part of a larger Open Society Project focused on trust in public institutions and punitive attitudes. The survey included questions regarding punitive attitudes, citizens’ experiences of victimisation and fear of crime, political ideology, trust in public institutions and neighbourhood/community cohesion. The target population were male and female adult residents (18 years-old or older) of urban and rural areas residents in Uruguay, carried out in 2018. The cases were selected using random sampling of cellular phones and the Computer-Assisted Telephone Interviewing (CATI) method. Cases were selected by weighting by telephone companies of the market and using quotas by sex, age and region (capital city vs. rest of the country) using national projections of the National Institute of Statistics of Uruguay (INE). Additionally, the sample was adjusted using weights based on the 2016 Index of Socioeconomic Status developed by the Centro de Investigaciones Economicas (CINVE). Cases were selected randomly until reaching pre-established quotas. The sample
size provided a sample error of 3.3% (95% confidence interval, CI). Participation in the survey was voluntary, and no reward was provided. Participants were informed about confidentiality issues and that they could abandon the interview at any time they wished.

The final sample included 465 females and 430 males aged 18–90 years (\(M = 44.15, SD = 17.65\)). A total of 78% of the sample were white, 10% mixed race and 6% black, with the remaining 6% Asian, indigenous or other. Less than 3% of the sample had not finished school, 10% had not finished high school, and 13% had obtained a university degree. Given the relevance that crime might have on fear of crime and on attitudes toward punishment (Hartnagel & Templeton, 2012), it is important to provide the context of criminality and insecurity of the sample. The present study was conducted in a country where violence has been systematically increasing. In the last three decades, there was almost a doubling of homicide rates per every 100,000 individuals (from 6.6 in 1990 to 12.8 in 2018; Galain et al., 2019; Ministry of Interior, 2019). Currently, Uruguay has the 4th highest homicide rate in South America after Venezuela, Brazil and Colombia, being one of the three top subregions of the world with the highest rates of homicide (United Nations Office on Drugs and Crime (UNODC), 2019).

5.2. Measures
We selected multiple items related to punitive attitudes mentioned in the literature that tapped into key dimensions relevant for evaluating the structure of this construct. The pool of items was taken from the most relevant and cited studies conducted by the top scholars on the field between 1980 and 2017 that included the most validated scales. These studies were systematically searched using Google scholar, Web of science and Scopus using the key words that included punishment, punitiveness, punitivity, punitive attitudes, penal attitudes, capital punishment attitudes, public opinion and punishment, public opinion and capital punishment, rehabilitation and public opinion. Thus, to adequately evaluate the dimensionality of punitive attitudes, we include a pool of 19 items that tapped into the severity of punishment, punitive reasons and progressive reasons, including multiple goals of punishment. Selection of dimensions and items involved a trade-off between following the most well-known and validated studies in the field and most relevant themes mentioned by the literature mentioned in Section 2.1. Particularly, we consider that evaluation of dimensionality requires including not only how harsh or intense punishment ought to be but also both types of reasons for supporting penal punishment, particularly given the well-known ambivalent views of citizens on these issues. The eight items focused on respondents’ beliefs about the severity of public responses to crimes and deviations, including questions about increasing penalties for adults or juveniles, the use of the death penalty, police or court abuse of human rights and even punishment outside the criminal justice system (e.g. universities). Six items tapped into punitive reasons for punishing offenders associated with goals such as retribution, deterrence and incapacitation. Finally, five items included more progressive reasons for punishing crimes associated with rehabilitation and restoration or reparation of harm. Each of the 19 items was measured using a Likert-type response that ranged from totally disagree (1) to totally agree (7) (Table 2).

5.3. Analytical plan
We start our analysis with a reliability analysis using Cronbach’s alpha for the global scale composed of the 19 items and for the three subdimensions (severity, punitive reasons, progressive reasons) and bivariate correlations between all items to explore general patterns and identify which items show poor functioning. Next, we use a battery of exploratory analyses to provide an initial evaluation of the structure of items and how unidimensional or multidimensional it is. We first conduct EFA
Table 2. Wording and distribution of the punitive items.

<table>
<thead>
<tr>
<th>Item</th>
<th>Label</th>
<th>Item from</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Make sentences more severe for all crimes’</td>
<td>sev1</td>
<td>Hogan et al. (2005)</td>
<td>895</td>
<td>5.65</td>
<td>1.90</td>
<td>7</td>
</tr>
<tr>
<td>‘A person convicted of murder should receive the death penalty’</td>
<td>sev2</td>
<td>Tyler and Boeckmann (1997)</td>
<td>895</td>
<td>3.72</td>
<td>2.48</td>
<td>4</td>
</tr>
<tr>
<td>‘Community alternatives (e.g. probation, community service, electronic monitoring) should be assigned more often’</td>
<td>sev3</td>
<td>Calaway et al. (2016); Melvin et al. (1985)</td>
<td>895</td>
<td>5.29</td>
<td>2.04</td>
<td>6</td>
</tr>
<tr>
<td>‘It is all right for the police to break the law in order to better control violent crimes’</td>
<td>sev4</td>
<td>Maguire and Johnson (2015)</td>
<td>895</td>
<td>4.54</td>
<td>2.54</td>
<td>5</td>
</tr>
<tr>
<td>‘In general, our courts have been more concerned with the rights of criminals than victims’</td>
<td>sev5</td>
<td>Viney et al. (1982)</td>
<td>895</td>
<td>5.58</td>
<td>1.99</td>
<td>7</td>
</tr>
<tr>
<td>‘Universities and military academies should have a policy of automatically dismissing students who are caught cheating on examinations’</td>
<td>sev6</td>
<td>Viney et al. (1982)</td>
<td>895</td>
<td>4.42</td>
<td>2.17</td>
<td>5</td>
</tr>
<tr>
<td>‘Juveniles who commit serious crimes should be treated like adults’</td>
<td>sev7</td>
<td>Mascini and Houtman (2006)</td>
<td>895</td>
<td>4.90</td>
<td>2.28</td>
<td>6</td>
</tr>
<tr>
<td>‘Make sentences more severe for juveniles who commit crimes’</td>
<td>sev8</td>
<td>Pickett et al. (2013)</td>
<td>895</td>
<td>5.39</td>
<td>2.07</td>
<td>6</td>
</tr>
<tr>
<td>‘Punishing criminals more harshly would reduce crime by setting an example and showing others that crime does not pay’</td>
<td>rpun1</td>
<td>Cullen et al. (1988)</td>
<td>895</td>
<td>5.37</td>
<td>2.08</td>
<td>6</td>
</tr>
<tr>
<td>‘Criminals deserve to be punished because they have harmed society’</td>
<td>rpun2</td>
<td>Cullen et al. (1985)</td>
<td>895</td>
<td>6.22</td>
<td>1.39</td>
<td>7</td>
</tr>
<tr>
<td>‘The amount of punishment that a criminal receives should be equal to the harm that the victim of the crime was forced to suffer’</td>
<td>rpun3</td>
<td>Cullen et al. (1988)</td>
<td>895</td>
<td>5.71</td>
<td>1.88</td>
<td>7</td>
</tr>
<tr>
<td>‘We should put criminals in jail so that innocent citizens will be protected from criminals who will victimise them’</td>
<td>rpun4</td>
<td>Cullen et al. (1985, 1988)</td>
<td>895</td>
<td>6.05</td>
<td>1.54</td>
<td>7</td>
</tr>
</tbody>
</table>

(Continued)
exploring the eigenvalues and scree plot to understand how many factors are observed. We also apply Velicer’s Minimum Average Partial (MAP) and parallel analysis to evaluate how robust initial results are regarding the construct’s dimensionality. Then, attention is turned to confirmatory analysis to evaluate whether CFA shows a similar or different picture in terms of the structure of the construct and the functioning of different items. Particularly, we compare the fit of the single construct alternative in relation to a multidimensional alternative that includes three correlated subdimensions and the second higher order model. Finally, we conduct bifactor analysis, both confirmatory and exploratory, to provide a more robust analysis of the structure of the construct.

6. Analysis

Initial analysis of Cronbach’s alpha shows that an overall reliability of the scale composed of the 19 items is .88, the severity subdimension
Table 3. Exploratory factor analysis.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Difference</th>
<th>Proportion</th>
<th>Cumulative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>6.626</td>
<td>5.392</td>
<td>.349</td>
<td>.349</td>
</tr>
<tr>
<td>Factor 2</td>
<td>1.234</td>
<td>0.029</td>
<td>.065</td>
<td>.414</td>
</tr>
<tr>
<td>Factor 3</td>
<td>1.204</td>
<td>0.219</td>
<td>.063</td>
<td>.477</td>
</tr>
<tr>
<td>Factor 4</td>
<td>0.986</td>
<td>0.069</td>
<td>.052</td>
<td>.529</td>
</tr>
<tr>
<td>Factor 5</td>
<td>0.916</td>
<td>0.04</td>
<td>.048</td>
<td>.577</td>
</tr>
<tr>
<td>Factor 6</td>
<td>0.877</td>
<td>0.022</td>
<td>.046</td>
<td>.623</td>
</tr>
<tr>
<td>Factor 7</td>
<td>0.854</td>
<td>0.08</td>
<td>.045</td>
<td>.668</td>
</tr>
<tr>
<td>Factor 8</td>
<td>0.775</td>
<td>0.047</td>
<td>.041</td>
<td>.709</td>
</tr>
<tr>
<td>Factor 9</td>
<td>0.728</td>
<td>0.06</td>
<td>.038</td>
<td>.747</td>
</tr>
<tr>
<td>Factor 10</td>
<td>0.668</td>
<td>0.05</td>
<td>.035</td>
<td>.783</td>
</tr>
<tr>
<td>Factor 11</td>
<td>0.617</td>
<td>0.012</td>
<td>.033</td>
<td>.815</td>
</tr>
<tr>
<td>Factor 12</td>
<td>0.605</td>
<td>0.096</td>
<td>.032</td>
<td>.847</td>
</tr>
<tr>
<td>Factor 13</td>
<td>0.509</td>
<td>0.013</td>
<td>.027</td>
<td>.874</td>
</tr>
<tr>
<td>Factor 14</td>
<td>0.496</td>
<td>0.021</td>
<td>.026</td>
<td>.9</td>
</tr>
<tr>
<td>Factor 15</td>
<td>0.475</td>
<td>0.051</td>
<td>.025</td>
<td>.925</td>
</tr>
<tr>
<td>Factor 16</td>
<td>0.423</td>
<td>0.043</td>
<td>.022</td>
<td>.947</td>
</tr>
<tr>
<td>Factor 17</td>
<td>0.38</td>
<td>0.034</td>
<td>.02</td>
<td>.967</td>
</tr>
<tr>
<td>Factor 18</td>
<td>0.347</td>
<td>0.067</td>
<td>.018</td>
<td>.985</td>
</tr>
<tr>
<td>Factor 19</td>
<td>0.28</td>
<td></td>
<td>.015</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Likelihood ratio (LR) test: independent vs. saturated: $\chi^2(55) = 5634.28$, Prob $> \chi^2 = 0.0000$; Kaiser-Meyer-Olkin (KMO) = .925; Determinant $> .061$. Own elaboration.

Figure 1. Correlation matrix.
is .74, punitive reasons subdimension is .78, and progressive reasons subdimension is .59 (the only one below the .7 accepted threshold). Three items show low item/retest correlation, and the overall reliability of the scale would improve if they were excluded (sev3, sev6, rpro5). These items show weaker correlations with the rest of the items of the scale (Figure 1). Bivariate correlations between the 19 items do not reveal a clear pattern with correlations of higher magnitude between items from the same subdimension (punitive reasons, progressive reasons or severity) in relation to correlations with items from other subdimensions (Figure 1).

EFA with no rotation shows that three eigenvalues have values greater than 1 (Table 3), and a three-factor solution would be appropriate according to Kaiser criterion (Kaiser, 1960). Yet, the difference in eigenvalue between the first and second factors (6.26 − 1.234 = 5.392) is greater than any other differences between subsequent factors. This is observed in the position of the elbow in the scree plot, which shows that the largest and most obvious break is between the first and second eigenvalues (Figure 2). Thus, it is reasonable to retain one factor and assume a one-factor dimension punitive scale. 3 Velicer’s MAP (Velicer, 1976) achieves a minimum of .01 retaining one factor, also supporting the unidimensionality hypothesis.

However, parallel analysis shows a different story. By taking into account the eigenvalues extracted from each factor for a set of randomised datasets with the same number of variables and cases as the initial data, we observe the number of factors that explain variance significantly differently from the variance explained by factors extracted from random data with the same characteristics (Howard, 2016). Results reject the unidimensional hypothesis and suggest a six-factor solution (Figure 2). Hayton et al. (2004) evaluate different estimation methods with respect to the number of factors in exploratory analysis, finding parallel analysis and Velicer MAP to be the most accurate. However, parallel analysis seems to tend to overestimate factors when it errs, and the opposite is true for Velicer MAP.

CFA was conducted to detect the presence of a unidimensional construct of punitiveness by examining the fit to the data of different models. Fit indexes show that the model with punitive attitudes as a single construct (Figure 3) is the one that fits the data worst: although the comparative fit index (CFI) and Tucker–Lewis Index (TLI) are above critical values of .95 (.982 and .979, respectively) and show good fit, the ratio of $\chi^2/df$ is 5.23, above the critical value of 5, and the value of the root mean square error of approximation (RMSEA) is .079, above the critical value of .06 (Table 4).

The fit of the model improves when we incorporate the assumption that punitive attitudes involve three subdimensions (severity, punitive reasons and progressive reasons). Both the correlated three-factor solution model (Figure 4) and the second higher order model (Figure 5) are equivalent and have equal fit: CFI and TLI show slightly better values of

---

3Analysis of individual items again shows that sev3, sev6, rpro5 load poorly on Factor 1 (loading more strongly to Factor 3). These results suggest that these items might be representing a different latent trait. They were eliminated from further analysis. Moreover, two additional items were eliminated: First, rpun5 was discarded due to its complex wording: it is difficult to understand and is based on a popular saying that is not necessarily widespread in the population. Item rpro4 was eliminated because of its problematic structure, which involved a double-barrelled question.
Figure 3. Global confirmatory factor analysis (CFA) model.

Figure 4. Three-dimensions confirmatory factor analysis (CFA) model.
.983 and .980, respectively. The ratio of $\chi^2/df$ is 5.06, still over the critical value, and likewise, the value of RMSEA is still over the critical value (.078; Table 4). However, the three-factor solution shows very high correlations between the three constructs, while the second-order model reveals a Heywood case between the progressive reasons construct and the general factor (g) with factor loadings higher than 1 and a negative error variance estimate. This result may indicate problems of model specification (Kolenikov & Bollen, 2012), particularly the impossibility of distinguishing the subdimensions from g (Bornovalova et al., 2020; Cucina & Byle, 2017).

Thus, we included a bifactor analysis to evaluate more robustly the unidimensionality of punitive attitudes (Dunn & McCray, 2020). We tested two models: one that included one general factor of punitive attitudes and three subordinate factors (severity, punitive reasons and progressive reasons), which did not converge; and one that

### Table 4. Fit measures.

<table>
<thead>
<tr>
<th>Models</th>
<th>$\chi^2$</th>
<th>$df$</th>
<th>CFI</th>
<th>RMSEA</th>
<th>SRMR</th>
<th>GFI</th>
<th>TLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global</td>
<td>507.255</td>
<td>77</td>
<td>.984</td>
<td>.079</td>
<td>.059</td>
<td>.989</td>
<td>.981</td>
</tr>
<tr>
<td>Three dimensions</td>
<td>473.785</td>
<td>74</td>
<td>.985</td>
<td>.078</td>
<td>.057</td>
<td>.99</td>
<td>.981</td>
</tr>
<tr>
<td>Second order</td>
<td>473.785</td>
<td>74</td>
<td>.985</td>
<td>.078</td>
<td>.057</td>
<td>.99</td>
<td>.981</td>
</tr>
<tr>
<td>Bifactor</td>
<td>267.197</td>
<td>63</td>
<td>.992</td>
<td>.06</td>
<td>.046</td>
<td>.994</td>
<td>.989</td>
</tr>
</tbody>
</table>

Note: CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual; GFI = goodness of fit index; TLI = Tucker–Lewis Index.

![Figure 5. Second-order confirmatory factor analysis (CFA) model.](image-url)
included one general factor of punitive attitudes and two subordinate factors (severity and reasons; Figure 6). This last model shows a better fit in relation to previous models: ratio of $\chi^2/df$ is 4.24 below the critical value; the value of RMSEA is not over the critical value (.06); the CFI and TLI also show improvement and are above the critical value (.992 and .989, respectively; Table 4). An additional argument for unidimensionality in this bifactor model is that the proportion of common variance across items explained by the general dimension (explained common variance, ECV) is .855 (above critical value .85). Also, when comparing the general and individual loadings of the bifactor model, most of them show stronger loadings toward the general model (except for sev7, sev8, rpun2 and rpro3). In terms of individual explained common variance (IECV), all items are above .5, and only three items are below .7 (rpun2, sev7 and sev8).

Additional reliability analysis of the scale provided support for the unidimensionality hypothesis. The omega coefficient was .85, and the hierarchical omega was .80, showing that most of the variance in the observed scores can be attributed to the general factor (Reise et al., 2013).

As a final robustness test, we conducted an exploratory bifactor analysis to evaluate the existence of other subdimensions above and beyond the general factor not considered in the theoretical model. We used an EFA approach based on the Schmid–Leiman (SL) transformation to calculate both the indirect effects of the general factor and the direct effects of the residualised group factors (Dombrowski et al., 2019). Our results (Figure 7) are consistent with the confirmatory bifactor analysis, with higher loadings on the general factor and a high hierarchical omega (.78), suggesting the presence of a unidimensional construct (Reise et al., 2013).
7. Discussion

Criminologists have increasingly considered public opinions about penal punishment an important topic in recent decades. However, little research has been conducted on carefully conceptualising and evaluating their psychometric structure. The scarce literature is divided by disagreement as to whether it is warranted to consider punitive attitudes a single concept or several different subdimensions.

Our results suggest that public opinion on penal punishment represents a single concept. While some of the initial analyses based on parallel analysis show evidence for different dimensions, exploratory factor and bifactor analysis suggest that punitive attitudes are a unidimensional scale. Overall, these results go in line with part of the literature (Aizpurúa, 2015; Armbrorst, 2017; Ortet-Fabregat & Pérez, 1992; Silver & Silver, 2017). Unlike Maguire & Johnson (2015), we do not believe that a unidimensional conceptualisation of punitive attitudes involves ignoring ‘meaningful complexity in how people think about criminal justice policy’ (Maguire & Johnson, 2015, p. 520) if it is adequately modelled with bifactor approaches that allow the evaluation of shared and unique elements in the construct.

However, the construct may be more complex and less consistent than expected (Armbrorst, 2017). First, some items were discarded, either because they were showing weak associations with the rest of the items of the scale or because of their problematic construction. Additionally, although our exploratory bifactor analysis suggests that most of the variance is explained by the general factor, our results still show unexpected subdimensional groupings in the residualised group factors. Thus, multidimensionality is still an open possibility, and future studies might challenge our results, showing more heterogeneity.

One relevant challenge in evaluating the structure of punitive attitudes is the type of questions used. Some of the items used in this study that are frequent in the literature have problematic reliability due to their double-barrelled nature. This is a significant issue when items oppose punitive to progressive attitudes (e.g. rehabilitation vs. punishment) since assuming them as opposites is problematic and misrepresent the more complex and contradictory perceptions of punishment (Cullen et al., 2000; Maguire & Johnson, 2015). Whether the hypothesis is that punitive and progressive attitudes have a negative, positive or even inconsistent relation, punitive and
progressive attitudes need to be measured in different items and, if possible, operationalised as different and specific reasons for punishment (e.g. retribution, specific deterrence, general deterrence, rehabilitation, etc.). Another challenge in the literature and in our study is the unidirectionality of most items, which might underestimate a negative relation between progressive and punitive attitudes (Baker et al., 2015) and thus generate biased results in the psychometric structure. Further difficulties in adequately modelling the heterogeneity of attitudes might be associated with the weak compatibility of items in terms of generality–specificity levels (Ajzen & Fishbein, 2005). Some items do not tap into attitudes toward penal punishment but instead into more general attitudes toward punishment. For example, when measuring retributive attitudes, some items tap specifically on punishing criminals ‘because they deserve it’ or ‘because they have harmed society’, while others tap on more general aspects related to the use of physical punishment by parents when children misbehave. Another potential source of bias could be the vague and heterogeneous definition and measurement of mental states implicit in items that mainly focus on the cognitive dimension. Developing a more explicit specification of emotional valences of items and including items that can adequately measure preferences clearly differentiated from mere acceptance (Aizpurúa, 2015) or even utility statements (Baker et al., 2015) could help to detect hidden heterogeneity in the punitive attitudes construct.

Criminological research on punitive attitudes could benefit from emulating psychology by incorporating a more rigorous validation of constructs. Although this is not common practice, there are some noteworthy exceptions in criminology where certain concepts have been more carefully examined: for example, the application of bifactor analysis to examine Gottfredson and Hirschi’s (1990) concept of self-control (see Bobbio & Arbach, 2020; Ward et al., 2015), or the utilisation of bifactor analysis, item response theory and measurement invariance to analyse Sampson and colleagues’ (1997) concept of collective efficacy (see Gerstner et al., 2019; Uchida et al., 2013). Future research on punitive attitudes should apply the aforementioned psychometric tests of the structure and validity of the construct in a more generalised manner (Armborst, 2017; Maguire & Johnson, 2015) because conclusions regarding its dimensionality might be strongly affected by the type of analysis. Additionally, future studies also need to address the cross-cultural challenge of measurement in criminology (Vazsonyi, 2003; Vazsonyi et al., 2021) and further evaluate measurement invariance of punitive attitudes constructs. We cannot assume that measures are reliable across very heterogeneous countries and regions (Armborst, 2017), particularly in those characterised by violence and more repressive, severe and weaker criminal justice institutions (Bergman & Fondevila, 2021).

Regarding the definition of punitive attitudes, borrowing the concept of attitudes from cognitive and social psychology is only a partial solution since the content and functioning of attitudes are also still under discussion in these disciplines (Ajzen & Fishbein, 2005; Albarracin & Shavitt, 2018). There is an agreement that attitudes are composed of cognitive, affective and behavioural components existing in the memory of individuals, which interact to give rise to the key aspect of attitudes – that is, the evaluation of an object (Albarracin & Johnson, 2018; Huskinson & Haddock, 2006). However, different components of attitudes can have different valences and even correlate negatively, leading to attitudinal ambivalence and reliability problems (Maio et al., 2004). In addition, the motivation and resources available to an individual when evaluating a certain object (e.g. offender, the criminal justice system, specific type of policies or laws, etc.) may activate different attitude recall mechanisms, leading to different outcomes, which change how punitive the
evaluation might be (Bohner & Dickel, 2011). Stating how different cognitive and emotional components of attitudes interact is highly complex, and conducting valid measurement of these elements is challenging given that implicit psychological processes are not always adequately captured by an explicit rating of self-report scales (Albarracin & Johnson, 2018; Carruthers, 2018) used by most criminological studies on punitive attitudes. Similar problems are observed when trying to establish the connection between attitudes with other mental states (beliefs, intentions or goals; Albarracin & Johnson, 2018) and its explanatory relationship with behaviours (Ajzen, 2012). This conceptual complexity in terms of the type of mental states is hardly ever captured by studies, mostly based on measures that tap only into some cognitive aspects of attitudes ignoring other aspects (e.g. emotional), and constitutes a significant challenge for future studies on punitiveness.

Having a clear idea of what punitive attitudes are and how they can be measured is a necessary precondition before trying to explain why some individuals are more punitive than others (Aizpurúa, 2015; Maguire & Johnson, 2015) and, particularly, what can be done to change those attitudes when they are based on mistaken assumptions about the criminal justice system. For example, punitive attitudes that assume that safety is associated with a significant increase in imprisonment are based on optimistic assumptions about deterrence. Showing offenders that crime does not pay through mass incarceration does not seem to significantly reduce recidivism, and has criminogenic effects (Cullen et al., 2011; Petrich et al., 2020) and even significant economic costs (Petersilia & Cullen, 2014). What role do these consequentialist considerations have in punitive attitudes in relation to the more retributive ones? How might individuals with punitive attitudes change their opinion or be convinced by more practical considerations of efficacy or cost efficiency? (Unnever & Cullen, 2010; Vuk et al., 2020). Unfortunately, research on the determinants of punitive attitudes, and specifically on how pragmatic considerations might play a significant role, is based on simplistic or problematic conceptualisations. If we want to identify the main drivers of punitive attitudes, and particularly those that are dynamic — that is, amenable to modification — we need to incorporate more adequate constructs as dependent variables in the explanatory models. Knowing the complexity of public attitudes toward punishment and how they can be modified is important for politicians and policymakers who sometimes feel pressured to promote inefficient and costly populist policies based on weak evidence of public opinion preferences (Garland, 2021). This is particularly relevant for low-income societies characterised by high levels of violent crime, limited resources and fragile and poorly organised criminal justice systems (Bergman, 2018; Maguire & Johnson, 2015).

The current study is not without some limitations. First, it was restricted to a representative sample of adult residents of urban and rural areas in Uruguay, and thus these results might not generalise to populations from other countries. Since most of the research evaluating the structure of punitive attitudes has been conducted in the United States and Europe with the exception of one study conducted in seven Caribbean countries (Maguire & Johnson, 2015), future research should corroborate if our findings are generalisable or associated with idiosyncratic characteristics of Latin-American countries. Specifically, research has shown that economic strains and high levels of victimisation, which characterise several cities of the region, are predictors of punitiveness (Fortete & Cesano, 2009; Lehmann et al., 2020; Singer et al., 2020). Thus, future studies should evaluate whether our psychometric results replicate in samples from less violent and more affluent contexts, where support for punitive and rehabilitation items might be different. Second, one important restriction is the content of items
used in this study and particularly that they did not allow heterogeneity to be distinguished in the dimensions of punitive and progressive reasons. For example, this study includes only two items that tap into retribution, only one that taps into specific deterrence and one on general deterrence. Likewise, under progressive reasons, only two items refer to restoration, and none are associated with re-entry programmes or social or early crime prevention programmes. Additionally, even when considering two items referring to the same reason, they might still yield different results. It remains for future studies to evaluate whether our results remain robust when considering more adequately all the different punishment reasons but also across various possible items. Third, most of the items used in this study follow the literature and focus on acceptance rather than on preference for punishing, let alone including more sophisticated measures of mental states such as emotions or sentiments. Future research should include items that not only tap on different types of reasons for punishment but also go beyond mere acceptance, allowing more meaningful relations between items to be explored and testing the dimensionality of public attitudes toward punishment in a more thorough way. Finally, the items included in our study are unidirectionally positive, which may introduce measurement bias associated with acquiescence or ‘yes-answering’. This implies a higher level of support for items irrespective of their specific punitive or progressive content (Pickett & Baker, 2014; Schuman & Presser, 1981). This tendency to respond positively to Likert scale items, influenced by social desirability or limited cognitive sophistication, can alter the relationships between items, leading to an overestimation of reliability scores for scales and the strength of associations between items in punitive studies (Pickett & Baker, 2014). This phenomenon is not confined to punitive studies but extends to the evaluation of attitudes in psychology, communication studies and political science (Hill & Roberts, 2023; Kuru & Pasek, 2016). Future research should replicate and assess the robustness of our psychometric results by incorporating batteries of bidirectional items.

**Ethical standards**

**Declaration of conflicts of interest**

Nicolás Trajtenberg has declared no conflicts of interest

Pablo Ezquerra has declared no conflicts of interest

Matthew Williams has declared no conflicts of interest

**Ethical approval**

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee [Cardiff University, ref. SREC/4283] and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent**

Informed consent was obtained from all individual participants included in the study.

**Acknowledgements**

Thanks to two anonymous reviewers and to the Reading Sessions in Quantitative Criminology (RESQUANT) group for their comments.

**Funding**

This work was supported by the Open Society Foundation Grant awarded to the project: ‘Punitiveness in Uruguay: how punitive are Uruguayans and what are the main reasons?’. The funders had no role in the study design, data collection and analysis, interpolation of these data, decision to publish or preparation of the manuscript.
References

ORCID
Nicolás Trajtenberg http://orcid.org/0000-0002-4451-3874


