



Sentencing, severity, and social norms: A rank-based model of contextual influence on judgments of crimes and punishments

Silvio Aldrovandi ^{a,*}, Alex M. Wood ^b, Gordon D.A. Brown ^a

^a Department of Psychology, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, United Kingdom

^b Behavioural Science Centre, Stirling Management School, University of Stirling, 3Y8 Cottrell Building, Stirling FK9 4LA, United Kingdom

ARTICLE INFO

Article history:

Received 11 May 2013

Received in revised form 8 September 2013

Accepted 17 September 2013

Available online 18 October 2013

PsycINFO codes:

2340

3040

Keywords:

Sentencing judgments

Crime seriousness

Social norms

Context effects

Decision by sampling

ABSTRACT

Context effects have been shown to bias lay people's evaluations of the severity of crimes and punishments. To investigate the cognitive mechanisms behind these effects, we develop and apply a rank-based social norms approach to judgments of perceived crime seriousness and sentence appropriateness. In Study 1, we find that (a) people believe on average that 84% of people illegally download software more than they do themselves and (b) their judged severity of, and concern about, their own illegal software downloading is predicted not by its amount but by how this amount is believed (typically inaccurately) to rank within a social comparison distribution. Studies 2 and 3 find that the judged appropriateness of a given sentence length is highly dependent on the length of other sentences available in the decision-making context: The same objective sentence was judged as approximately four times stricter when it was the second longest sentence being considered than when it was the fifth longest. It is concluded that the same mechanisms that are used to judge the magnitude of psychophysical stimuli bias judgments about legal matters.

© 2013 Published by Elsevier B.V.

1. Introduction

In an attempt to make application of the law as fair and unbiased as possible, various attempts at standardization have been made. Sentencing guidelines are an example of this approach, as they constrain the array of potential sentences to be handed out given the circumstances of the case under consideration (e.g., Ruback & Wroblewski, 2001). Similarly, several indices of the seriousness of crimes have been compiled (e.g., Sellin & Wolfgang, 1964). However, the success of these attempts may be undermined by strategies adopted by legal professionals when reaching a decision, such as the adoption of 'mental shortcuts' and the neglect of relevant information (e.g., Dhimi, 2003; Dhimi & Ayton, 2001; von Helversen & Rieskamp, 2009).

In addition, lay people's evaluations are widely variable and influenced by contextual and extra-evidential factors (e.g., Peer & Gamliel, in press; Stylianou, 2003). Such influences may undermine public consensus about the seriousness of crimes (e.g., Durham, 1993; Gerbasi, Zuckerman, & Reis, 1977; Hoffman & Hardyman, 1986; Kwan, Chiu, Ip, & Kwan, 2002; Roberts, 1992; Rossi, Simpson, & Miller, 1985).

Lay people's evaluations of legal matters influence law-related policies and decisions (e.g., Roberts, 2003; Roberts & Doob, 1989; Roberts & Stalans, 1997; Stalans & Diamond, 1990). For example, citizens in the UK can refer specific sentences, if deemed too lenient, to the Attorney General, who can review the sentence and increase the punishment. Context-induced biases could be seen to threaten the basis for this process, as it has been shown that people tend to perceive sentencing as too lenient, although this bias is fuelled by misconceptions and misinformation about the current sentencing guidelines, the severity of penalties handed out and of the crimes committed (e.g., Roberts, 2003; Stalans, 1993).

In order to determine whether the same cognitive mechanisms may be responsible for a range of contextual effects that bias lay people's evaluations of the law, we put forward—and systematically test—a relative judgment model of legal evaluations. We hypothesize that judgments about the seriousness of a crime or the appropriateness of a sentence are made in comparison to relevant information that is retrieved from memory (e.g., inaccurate beliefs about the frequency of crimes) and available in the decision-making context (e.g., sentences handed out for similar crimes). Consequently, the same sentence (or the same crime) might attract very different evaluations as a result of the influence of contextual factors. Before describing the model of how context affects legal evaluations, we first briefly discuss evidence for relative judgment within the two above-mentioned legal domains: The perceived seriousness of crimes and fairness of sentences.

* Corresponding author. Tel.: +44 24765 28387.

E-mail addresses: S.Aldrovandi@warwick.ac.uk (S. Aldrovandi), Alex.Wood@stir.ac.uk (A.M. Wood), G.D.A.Brown@warwick.ac.uk (G.D.A. Brown).

1.1. Sentencing appropriateness, crime seriousness and social norms

Previous research has showed that, despite the federal guidelines, sentencing varies to some degree between courts due both to the characteristics of the individual decision-makers (the judges) and to the social interaction within courtrooms (e.g., Bushway, Owens, & Piehl, 2012; Eisenstein & Jacob, 1977; Johnson, 2006). In the present paper we focus on the cognitive processes that underpin context effects rather than the sociological aspects of sentencing (e.g., Hamilton & Rytina, 1980; Stylianou, 2003), and draw on evidence that suggests that people rely on relative comparisons when evaluating the appropriateness of sentences. For instance, people's punitive attitudes might be rooted in comparisons made against inaccurate impressions they hold about crimes: People think that most burglaries involve the use of weapons, property damage, and harm—although in reality this is rarely the case (e.g., Roberts & Doob, 1989; Stalans, 1993; Stalans & Diamond, 1990). Thus, a key motivation for our research is that people tend not to make absolute judgments, but instead rely on relative comparisons (e.g., Stewart, Brown, & Chater, 2005).

Relative comparisons are also at the core of socially-oriented accounts of dishonest behavior and consequent worry about it. For example, social norms (e.g., Campbell, 1964; Henrich et al., 2001) can be regarded as the rules and standards that are shared by members of a community. These norms guide behavior of the individuals without the need for higher intervention such as law enforcement (e.g., Mazar & Ariely, 2006; Mazar, On, & Ariely, 2008): For instance, people are more likely to pay taxes if they believe that others (especially their own friends) do so (Posner, 2000).

Thus, social norms theories can be interpreted as relative approaches: Judgments about the seriousness of a crime are determined by comparisons relative to beliefs about the social context. For instance, the perceived seriousness of (and decisions about) not paying taxes and littering depends on the beliefs about how many other people commit such crime and negligent behavior (e.g., Cialdini, Reno, & Kallgren, 1990; Orviska & Hudson, 2002; Posner, 2000; Traxler, 2010; Yankelovich, 1984). However, the precise nature of the comparisons made to hypothesized internalized standards has not been extensively explored. Thus, we examine in more detail the relationship between the frequency of an individual's own dishonest behavior and their beliefs about the frequency of other people's dishonest behavior (the internalized standards). We specify the cognitive processes underlying these comparisons, the outcome of which determines the perceived seriousness of law infringements and of crimes in general.

1.2. A relative judgment model

We hypothesize that the cognitive processes that underlie lay people's judgments about the seriousness of crimes and the appropriateness of sentences are the same as those underpinning judgments about other psychophysical, economic and social stimuli. Several theoretical accounts have been developed to explain quantitatively how judgments are influenced by comparison context; in this paper, we test the predictions of these theories as applied to judgments of sentence severity and crime seriousness. Adaptation Level Theory (ALT; Helson, 1947, 1948) proposes that people's evaluations of stimuli are a function of all previously experienced relevant stimuli. As people experience reality, they form an internalized 'reference level' for a particular outcome ('the adaptation level'); any incoming stimuli will be then be judged against the reference level. For example, a person might believe that, on average, people drive at 80 miles/h on motorways in the UK (i.e., 10 miles/h over the speed limit). When driving, this person will therefore consider her driving of 5 miles/h over the speed limit as a relatively minor infraction, and its seriousness to be low. Similarly, according to ALT, a recent sentence handed out for a given crime would be evaluated as lenient, appropriate or strict depending on how much it differs from the mean of sentences encountered for similar crimes.

However, the effects of other contextual indicators have been observed in a series of evaluation tasks, and suggest that ALT may not be able to account for a series of phenomena. First, the *range* of the contextual stimuli can influence the subjective evaluation of a stimulus (e.g., Janiszewski & Lichtenstein, 1999; Volkmann, 1951). That is, the subjective value of a stimulus depends on where it falls within the set of values of the reference distribution—i.e., how far it deviates from the smallest and largest values in the set (the *range* principle). Thus, the same stimulus might be evaluated differently depending on prior beliefs. For instance, the appraisal of the fairness of a sentence of 3 years for a crime of robbery might be determined by beliefs about the typical maximum and minimum sentence lengths handed out for robberies in general. A person who believes that the sentences generally handed out for robberies range from 1 to 5 years will be inclined to regard the 3-year sentence as more severe than will a person who believes that the sentences for robbery range from 1 to 9 years of imprisonment.

Second, *rank* effects have been observed, whereby the subjective magnitude of a stimulus is determined by its ranked position within the contextual stimuli. The Decision by Sampling model (DbS; Stewart, Chater, & Brown, 2006) offers a description of the psychological processes underpinning rank effects, although DbS has not previously been applied in legal contexts. Applied to the present context, DbS hypothesizes that people, when facing with a task such as evaluating the fairness of a given sentence, retrieve from memory a sample of sentences handed out in similar situations; they also sample instances that are relevant to the situation at hand and that are available in the decision-making context. The subjective value of the sentence under consideration (e.g., its fairness) is determined by a series of binary ordinal comparison between the sentence length itself and the lengths of the sentences (a) retrieved from memory and (b) available in the decision-making context. The perceived fairness of the sentence is directly determined by its relative ranked position within the sample. For instance, if someone has to evaluate the fairness of a sentence of 2 years in prison for a case of burglary, she might retrieve from memory a few other sentence lengths that she has recently encountered; as an example, six of these might be shorter than 2 years, while two are actually longer. The relative rank value of the sentence of 2 years would therefore be calculated according to that retrieved sample: (number ranked lower)/(sample size) = .75. As long as different samples are retrieved from memory (e.g., because of differences in the use of the media; Gebotys, Roberts, & DasGupta, 1988; Graber, 1980; Roberts & Doob, 1989; Stalans, 1993), the same verdict might be regarded rather differently. We test this in Study 1.

Range Frequency Theory (RFT; Parducci, 1965) offers a descriptive model of rank and range effects. Specifically, assume an ordered set of n contextual items $[x_1, x_2, \dots, x_i, \dots, x_n]$. Then, if M_i is the subjective psychological magnitude of x_i ,

$$M_i = wR_i + (1-w)F_i \quad (1)$$

where R_i is the range value of stimulus x_i :

$$R_i = \frac{x_i - x_1}{x_n - x_1} \quad (2)$$

and F_i is the frequency value, or relative ranked ordinal position, of the item i in the ordered set:

$$F_i = \frac{i-1}{n-1} \quad (3)$$

Eq. (1) embodies the claim of RFT that the evaluation of a stimulus is a compromise of both rank (Eq. 3) and range (Eq. 2) principles, with w being a weighting parameter.

Rank and range principles were initially validated in the domain of psychophysics (Parducci, Calfee, Marshall, & Davidson, 1960; Parducci & Perrett, 1971) where people performed such tasks as evaluating the loudness of sounds. The theory later received considerable experimental

support in various cognitive and social domains, including satisfaction with own salary (Brown, Gardner, Oswald, & Qian, 2008; Hagerty, 2000; Parducci, 1995) and life in general (Boyce, Brown, & Moore, 2010; Smith, Diener, & Wedell, 1989); price perception (Niedrich, Weathers, Hill, & Bell, 2009); attitudes towards the riskiness of alcohol consumption (Wood, Brown, & Maltby, 2012); the effects of income on psychopathology (Wood, Boyce, Moore, & Brown, 2012); emotional reactions to events (Wood, Brown, & Maltby, 2011); evaluation of pain (Watkinson, Wood, Lloyd, & Brown, 2013) and of anxiety and depression symptom severity (Melrose, Brown, & Wood, 2013); beliefs about the health benefits of physical exercise (Maltby, Wood, Vlaev, Taylor, & Brown, 2012); judgments about personality (Wood, Brown, Maltby, & Watkinson, 2012) and the fairness of allocation of salaries and taxes (Mellers, 1982, 1986).

1.3. The current studies

In the present series of studies, we test the predictions of different cognitive models of decision-making on lay people's judgments about the seriousness of crimes (Study 1) and the appropriateness of sentencing (Studies 2 and 3).

Determining which cognitive theory best describes judgments about legal matters also has practical implications. We illustrate this point by focusing on judgments about the seriousness of a crime. We use downloading copyrighted material over the internet as an example of law infringement. If people evaluate the seriousness of their own law infringements solely in terms the frequency with which they commit them (as predicted by absolute approaches), then (a) contextual information will not impact their concern, which (b) will be determined by the frequency with which people infringe the law. For instance, as downloading copyrighted materials over the internet is on the rise, an absolute approach would predict an associated increase in level of concern—as the latter is directly determined by the extent of infringing the law.

Conversely, Adaptation Level Theory (Helson, 1947) would predict relatively unchanged levels of concern, despite the higher frequency of illegal downloading. As the increase in illegal downloading will move the reference point (adaptation level) upwards, the relationship between own illegal downloading and other people's average illegal downloading might not change (possibly after a period of adaptation). The same prediction can be made by the rank principle (Parducci, 1965; Stewart et al., 2006), according to which it would in contrast be the ranked position (for illegal downloading) within the reference distribution of illegal downloaders that does not change as overall levels increase.

The range principle can paradoxically predict a decrease in concern, as it states that subjective evaluations of a stimulus (i.e., concern about own illegal downloading) depend also on how such a stimulus differ from the extremes of the distribution (i.e., those who illegally download the least and the heaviest downloaders). It is therefore important to consider people's beliefs about the uniformity of the increase of illegal downloading across the distribution of illegal downloaders. If a person believes that only a few people increased their downloading (that is, the distribution of illegal downloading become more positively skewed), then the concern about her habit of illegally downloading material over the internet might decrease, as the range value of her own illegal downloading decreases (her downloading habit has become further away from that of the heaviest downloaders).

In Study 1, we explore the nature of the relative comparisons people make when evaluating the seriousness of the law infringement they themselves commit. In Studies 2 and 3 we tested the rank and range principles experimentally, through examining their influence on judgments of fairness of sentences.

2. Study 1

In this study we measured participants' attitudes towards downloading copyrighted material from the internet. We chose this particular

law infringement—instead of more severe crimes sentenced in courts—mainly because, due to its prevalence, participants are more likely to admit to it.

Previous research has shown that people perceive the seriousness of their law infringements depending on how they perceive social norms (standards) about them (e.g., Cialdini et al., 1990). However, no formal account has been proposed to quantitatively describe the comparisons people make to the internal standards. Thus, we tested participants' beliefs about the frequency of other people's law infringements (i.e., the amount of illegal downloading that other people do). We elicited from each participant their beliefs about the entire frequency distribution of illegal downloading done in the UK. We aim to test the predictions of different models of judgments about the law, specifically (1) the absolute approach, whereby perceived seriousness of law infringements is directly related to the extent to which the law is infringed, (2) ALT (Helson, 1947) which predicts that concern about breaking the law depends on how much one's rate of dishonest behavior exceeds or falls below some measure of the central tendency of other people's infringements and (3) the relative rank account, where ranked position predicts attitudes towards the infringements.

2.1. Method

2.1.1. Participants

A total of 135 participants (87 females) from a university in the UK volunteered to take part in this study. Ages ranged from 18 to 56 years ($M = 20.89$, $SD = 4.68$). Participants were included in the study if they downloaded copyrighted material from the internet. Those who were students were enrolled in a series of undergraduate courses, with psychology being the most prevalent (68.10%). The majority of participants were of White ethnic background (54.80%), while a minority was Chinese (15.60%) or Indian (12.60%).

2.1.2. Design and procedure

Participants were tested individually. The questionnaire, which took approximately 10 min to complete, had the following general introduction: "Copyright is a form of protection provided by the law to the authors of original works of authorship, including musical and artistic works. Copyright gives the creator of the work the exclusive right to make copies of it. Over the internet, people infringe copyright laws when they download copyrighted material—for instance music, video, films, software and videogames. The following questions are on your opinion about the total amount of copyrighted material which is illegally downloaded in UK from any source, including peer-to-peer filesharing. Please note that each song counts as 1, as each video, film, software program or videogame does."

The next section included 11 questions of a subjective probability elicitation task, which elicited participants' beliefs about the distribution of other people's downloading habits. There are different ways to elicit probability distributions (e.g., Lewandowsky, Griffiths, & Kalish, 2009; Manski, 2004); here we simply asked students to estimate different percentiles points of the distribution (e.g., Melrose et al., 2013; Wood et al., 2012). We used 11 questions phrased as follows: "The top $x\%$ most prolific (heavy) downloaders of copyrighted material download more than how many songs, videos, software, films and videogames per month typically?", where x [respectively] had values of [1, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 99]. The order of the questions was counterbalanced between subjects: Half of the participants estimated the downloading done by the $x\%$ most prolific downloader (descending order), while the remaining half estimated the downloading done by the $x\%$ least prolific (ascending order). A minority of participants responded incoherently in the subjective probability elicitation task, in that their estimates of the relevant percentiles did not increase (or decrease, in the descending order condition) with percentile. We therefore adopted a data exclusion condition: a participant's data were excluded if the Kendall's τ coefficient between stimuli (i.e. the 11 percentile points)

and responses (i.e. the participant's estimates of downloading for each of the 11 percentile points) was $<.50$. This process led to the removal of data from 23 participants (17.04% of the total); results do not change, either qualitatively or quantitatively, if these participants are included in the analyses.

Having elicited participants' beliefs about the social distribution of UK illegal downloading we asked participants to indicate how many copyrighted songs, videos, software, films and videogames (each counting as 1 unit) they themselves downloaded in a typical month. To verify the reliability of participants' responses about their own frequency of illegal downloading, we collected an additional measurement of participants' downloading habits: we later asked participants to indicate how many items they illegally downloaded each week, to enable assessment of participants' consistency in reporting their downloading habits.

We then measured attitudes towards law infringements by prompting participants' responses on two items. First, participants assessed the seriousness of their downloading habits by answering the question "When downloading copyrighted material, how seriously do you think you are infringing the law?" on a 1 ('Not serious at all') to 7 ('Extremely serious') point scale. Next, participants rated their worry as a consequence of their downloading by answering the question "How worried are you about being caught downloading copyrighted material?" on a 1 ('Not worried at all') to 7 ('Extremely worried') scale. These two items were chosen as they measure different aspects of concern as a result of infringing the law. The first item measured the wrongfulness of an offense (e.g., O'Connell & Whelan, 1996; Warr, 1989), while the second measured concern due to the perceived likelihood of getting caught and receiving punishment.

Next, we measured participants' entertainment habits in general, in order to control for participants' propensity towards listening to music, watching film etc., in general. We asked participants to indicate, for an average month, how many (a) music CDs they bought (b) songs they download from authorized websites (c) live concerts they went to (d) times they went to the cinema. We then added the scores of these four variables in order to generate a variable indicative of entertainment habits.

To control for participants' computer literacy, we collected participants' subjective assessments by asking them to assess their computer skills using a 7-point scale ranging from 1 ("Basic") to 7 ("Advanced"), where 4 represented an "Intermediate" literacy.

2.1.3. Statistical analyses

In order to calculate the rank position of each student within her inferred distribution of student illegal downloading, we estimated each student's cumulative distribution function to the answers to the 11 questions of the subjective probability elicitation task. We chose either a lognormal function or a linear function for each participant, depending on which fitted best. Example cumulative distribution functions are shown in Fig. 1 below. We then computed the mean of the distribution ('Subjective mean') to test the predictions of ALT (Helson, 1947). Finally, we computed the relative rank position of each student within her own inferred distribution of downloading ('Subjective rank').

To analyze participants' responses we used ordinal regression (using the polytomous universal model, PLUM). We first examined whether the order of the 11 questions eliciting the inferred distribution of student debt (i.e. ascending vs. descending orders) had an effect on any of the main analyses reported. As question order did not enter as a significant predictor in any of the analyses, data were collapsed across the two orders.

2.2. Results and brief discussion

We first looked at the consistency of participants' responses about their downloading habits. The two estimates of downloading habits (i.e. monthly and weekly) correlated highly, $r = .92$, $p < .001$. We

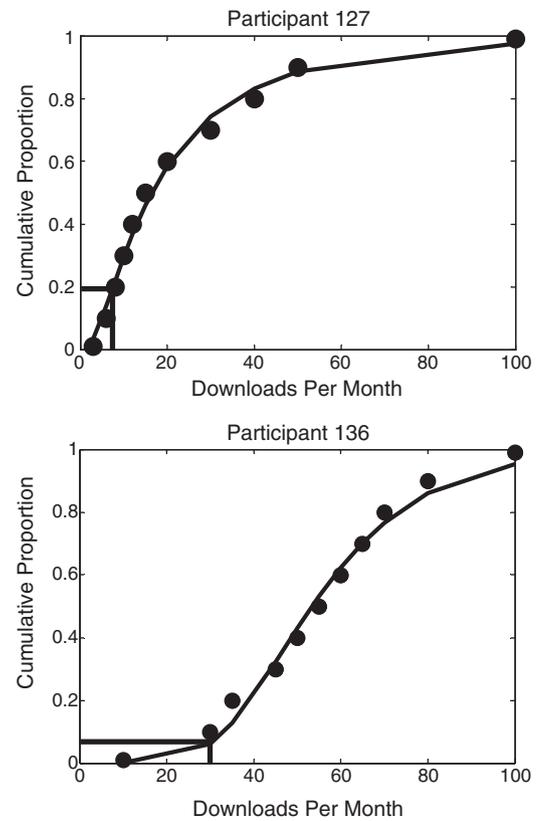


Fig. 1. The cumulative distribution (filled circles) along with a best-fit cumulative density function (solid lines) elicited from participants 127 (top panel) and 136 (bottom panel) in Study 1. Vertical lines indicate own downloading, while the horizontal lines represent the inferred rank position.

transformed the participants' weekly downloading frequency into a second monthly estimate (by multiplying it by 4.3, the number of weeks in a month). The difference between the two monthly estimates was not significantly different from zero across participants, $M = 0.15$, $SD = 17.59$, $d = 0.01$, although one participant was excluded from the analyses as the difference between her two estimates was 4.21 SDs larger than the mean; all other participants' z-scores ranged from -1.94 to 1.69 .

Table 1 displays data concerning participants' beliefs about the occurrence of illegal downloading in the UK. Participants displayed little agreement on the frequency of illegal downloading. For instance, for the median percentile of the inferred distribution, the interquartile range (IQR) of downloaded items estimates ranged from 33 to 300. Great variability was also detected for the 10th (IQR = [5, 50]) and 90th percentiles (IQR = [93, 1000]). Participants also largely underestimated their own rank position within the distribution of downloaders in the UK ($M = .16$, $SD = .14$; IQR = [.06, .22])—although the rank values ranged from .01 to .94. Similarly, by fitting each individual downloading amount within the distribution of

Table 1
Downloading estimates (in downloaded units) provided in the subjective probability elicitation task for low, medium and high percentile points.

Percentile points	Median	IQR	
		LL	UL
10th percentile	15	5	50
50th percentile	70	33	300
90th percentile	400	93	1000

Note. IQR = Interquartile Range; LL = lower limit; UL = upper limit.

Table 2
Regression coefficients for the analyses of Study 1.

	Estimate	SE	Wald	p
<i>Seriousness</i>				
Subjective rank	3.94	1.55	6.43	.011*
Downloads per month	<0.01	0.01	0.07	.787
Subjective mean	<0.01	<0.01	0.63	.429
Gender	−0.62	0.42	2.14	.143
Age	0.03	0.05	0.25	.619
Ethnicity	1.00	0.41	5.97	.015*
Entertainment	0.01	0.02	0.60	.439
Computer literacy	0.24	0.15	2.79	.095
<i>Worry</i>				
Subjective rank	4.05	1.59	6.47	.010*
Downloads per month	0.01	<0.01	3.07	.080
Subjective mean	<0.01	<0.01	0.22	.641
Gender	−0.28	0.43	0.43	.514
Age	0.04	0.05	0.73	.394
Ethnicity	0.94	0.42	5.06	.024*
Entertainment	−0.03	0.02	2.50	.114
Computer literacy	−0.10	0.15	0.43	.511

Note. *Significant at 5% level.

downloading done by the present sample, we were able to determine that the average rank underestimation was 40%; indeed, only 5 participants in total overestimated their own rank position (by an average of 7%). Overall, participants reported that they downloaded on average 20.06 items a month ($SD = 29.04$; $IQR = [5, 20]$).

In general, attitudes towards illegal downloading were rather permissive; the average seriousness rating was 2.99 (on a 1 to 7 scale; $SD = 1.31$), while the average score on the item measuring worry was 2.23 (on a 1 to 7 scale; $SD = 1.37$).

The large variation in beliefs about other people's frequency of illegal downloading can be exemplified by previewing the results from two participants. Fig. 1 shows the best-fit cumulative density functions (solid lines) to the beliefs about the number of items that are downloaded illegally by others each month (full circles), for participants 127 (top panel) and 136 (bottom panel). For instance, participant 127 believed that 70% of people download around 30 items each month, while participant 136 thought this value to be around 65. Participant 127 reported that she downloaded around 8 items each month (the vertical line in the figure); her inferred rank is around .20, which means that she believed that approximately 20% of other students downloaded less than she did. In contrast, participant 136's inferred rank was only .08, although her downloading of 30 items a month was almost four times as many as for the other participant. Consistently with this pattern, the reported concern about own illegal downloading was higher for participant 127 than for participant 136—although the latter actually downloaded more often than the former. This finding suggests that concern about own dishonest behavior is driven not by its frequency, but rather by the beliefs about how such frequency ranks among others.

Table 2 above shows the regression coefficients for the analyses of the perceived seriousness of own downloading habits and concern about them.¹ We tested against each other the predictors associated with the relative rank approach (e.g., Parducci, 1965; Stewart et al., 2006; 'Subjective rank'), ALT (Helson, 1947; 'Subjective mean'), and an absolute approach (i.e. the more one downloads, the more concerned

a participant will be; 'Downloads per month'). In addition, we entered the following predictors: (a) gender (1 = Male, 2 = Female), (b) age (c) ethnicity (1 = White, $n = 67$; 2 = Non-white, $n = 45$), (d) entertainment (the higher the scores, the more often one goes to the cinema, buys CDs etc.) and finally (e) computer literacy (the higher the scores, the higher the self-assessed computer literacy).

As it can be seen in Table 2, rank position significantly predicted both dependent variables. Strikingly, no significant contribution was added by the reported amount of downloading done, nor did it matter how one's downloading fared compared to the (inferred) UK average downloader. Thus no empirical support was obtained in favor of ALT (Helson, 1947), at least with adaptation level operationalized as the mean. Participants' evaluations about own downloading habit were best predicted by where they thought they ranked compared to other downloaders, as predicted by the rank principle of DbS (Stewart et al., 2006).

Overall, the results strongly suggest that students evaluate their own law infringements in relative terms. Rather than comparing themselves to some internalized standard (e.g., the adaptation level), or simply considering their own behavior irrespective of that of other people, participants' evaluations were driven by where they thought they ranked within what they believed to be distribution of illegal downloaders in the UK.

3. Study 2

In Study 1 we observed strong associations between the rank position of own crime frequency and its subjective seriousness. The study thus showed that the rank principle drives the perception of crime seriousness; a different context—different beliefs about crime prevalence—leads to a different perception of the very same crime.

In Study 2 we extend these findings to the evaluations of sentencing. Moreover, we tested the causality of the rank principle by manipulating the rank position of a given sentence length relative to other sentence lengths available in the decision-making context. In doing so, we test the rank principle. The aim of this study was to investigate whether judgments about the appropriateness of a given sentence is influenced by the immediate decision-making context in the same manner as the perception of crime seriousness is. Participants were asked to evaluate the appropriateness of a set of 11 different sentence lengths given on different occasions for the same type of crime. We manipulated the distribution of the sentence lengths in order to test the predictions of the rank principle.

To increase the generalizability of the results, two types of crime scenarios were selected: fraud and sexual offense. These crimes greatly differ in their nature and usually attract different sentence lengths (fraud attracted an average sentence length of 12 months in the period 2006–2008 in England and Wales, while the average sentence length was 45 months for sexual offenses; Ministry of Justice, 2010); they also differ in nature from the crime investigated in Study 1, and this will allow us to extend its findings to more serious crime scenarios. At the same time, we intentionally provided little detail about other circumstantial factors (e.g., characteristics of the victim and of the perpetrator, courtroom community type, rural/urban county, plea bargaining etc.), as these have already been shown to influence such evaluations (e.g., Stylianou, 2003).

3.1. Method

3.1.1. Participants

A total of 81 participants (55 females) from a university in the UK either volunteered or were rewarded with course credit to take part in this study. Age ranged from 18 to 77 years ($M = 25.53$, $SD = 12.99$). The majority of participants were of White ethnic background (55.56%); other ethnicities included Indian (6.17%) and Chinese (4.94%) participants.

¹ As some of the variables were positively skewed, we repeated the analyses on transformed data in order to reduce the skew (by both logarithmic and square root transformations). The same results were observed when these transformations were applied. For instance, when the data were logarithmically transformed, subjective rank ($B = 9.84$, $SE = 4.73$, $Wald = 4.33$, $p = .037$ and $B = 12.59$, $SE = 4.87$, $Wald = 6.67$, $p = .010$) and ethnicity ($B = 0.84$, $SE = 0.41$, $Wald = 4.17$, $p = .041$ and $B = 1.24$, $SE = 0.43$, $Wald = 8.46$, $p = .004$) were the only significant predictors for the seriousness and worry subjective ratings, respectively. None of the other variables were significant predictors for neither of the two dependent variables (all $ps > .07$).

3.1.2. Design and procedure

Participants filled in a 2-page questionnaire, one for each crime scenario. One each page, 11 sentence lengths (in months) for a given crime were presented. For each crime scenario, participants were told that the 11 values represented custodial sentences (in months) that were imposed on a perpetrator; that is, they represented how long s/he was sentenced to held in custody—i.e. sent to jail. Participants were asked to imagine that the crime was of moderate seriousness and that it was committed by a first time offender. In order to describe the two different crime scenarios, we changed the introductory text. For fraud, the instructions were as follows: “Fraud may be defined as the use of deception with the intention of obtaining an advantage, avoiding an obligation or causing a loss to another party. Acts of fraud may include bribery, forgery, and corruption.” For the sexual offense crime, the instructions read: “Sexual crimes involve the conduct of a sexual or indecent nature towards another person that is accompanied by actual or threatened physical force or that induces fear, shame, or mental suffering.” Below the instructions there was a table listing the 11 sentence lengths; for each, participants had to rate how appropriate they thought each sentence length was on a scale ranging from –3 (“much too lenient”) to +3 (“much too strict”), where 0 represented an “appropriate” sentence. As in Study 1, we removed from the analyses data from 10 participants who responded erratically (12.35% of the total), as either (a) the Kendall’s τ coefficient between stimuli (i.e. the sentence lengths presented) and responses (i.e. participants’ appropriateness ratings) was $<.50$ or (b) the response range for their appropriateness ratings was <1.00 . Data removal did not compromise the balance of the design; moreover, the results again did not change when all participants were included in the analyses.

The order of the presentation of the 2 crime severities—low (fraud) and high (sexual offense)—was counterbalanced between participants. The order in which the 11 sentences were presented to each participant was manipulated between-subjects and counterbalanced across crime scenarios. In the ascending order condition, the first of the 11 amounts presented was the shortest, while the opposite was true for the descending order condition.

We manipulated the distribution of the sentence lengths between participants: 41 participants were presented with a *bimodal* distribution, while the remainder attended to a *unimodal* distribution (see Table 3). Only five sentence lengths are in common between the two distribution types (the five ‘common points’).

In the sexual offense scenario, the shortest sentence length was the same in both distributions (12 months; common point 1). This length also has the same rank position within both distributions (i.e. its rank is equal to 1), hence no differences in participants’ subjective ratings are predicted by either the absolute or the relative accounts of

judgments. The second common point is 30 months; in the unimodal distribution it ranks as the 2nd shortest (i.e. rank = 2), while it ranks as the 5th in the bimodal distribution (rank = 5). If rank determines participants’ appropriateness ratings, 30 months should be rated as more lenient in the unimodal distribution than in the bimodal distribution, despite the actual length of the sentence being the same in both cases. Also, 30 months has the same distance from the distributions’ means (i.e. it is 15 months shorter), hence any difference cannot be readily accounted for by ALT (Helson, 1947), assuming that the reference level can be operationalized as the mean. The distribution mean (45 months) is the third common point; both its length and its rank position (rank = 6) are the same in both distributions; hence, as for the first common point, no differences in participants’ evaluations are expected according to any of the cognitive theories. Conversely, 60 months (common point 4) ranks lower in the bimodal distribution (rank = 7) than in the unimodal distribution (rank = 10)—hence we expect it to be rated as a stricter sentence in the unimodal distribution. As for the 2nd common point, common point 4 has the same distance from the distributions’ means as in both cases it is 15 months longer; hence, ALT would predict no differences. The longest sentence in the set (78 months; rank = 11) is the 5th common point; as for the first and third common points, no differences in subjective ratings are expected between distributions.

The design was a therefore a 2×2 mixed design. Distribution (unimodal vs. bimodal) was a between-subjects factor, while crime severity (low being fraud and high being sexual offense) was manipulated within subjects.

3.2. Results and brief discussion

We compared participants’ responses in the unimodal and bimodal distributions for both crime severities. Fig. 2 presents participants’ mean responses for the 5 sentence lengths which were presented in both conditions (common points). For the 1st, 3rd and 5th common points appropriateness ratings were comparable across crime scenarios; this finding was expected as each sentence had the same length and rank position within each distribution type. In line with the rank principle, the sentence length corresponding to common point 2 was rated as stricter in the bimodal distribution—where it ranked as 5th—compared to the unimodal distribution, where its rank was 2. As predicted, common point 4 attracted higher responses in the unimodal distribution (rank = 10) than in the bimodal distribution (rank = 7). This pattern of results was similar for both crime scenarios.

A 5 (within: common point) \times 2 (within: crime severity) \times 2 (between: distribution) mixed ANOVA confirmed the observations above. There was a significant main effect of point, as longer sentences attracted stricter ratings, $F(4, 252) = 724.82, p < .001, \eta^2_p = .92$. There was also the predicted interaction between distribution and common point, $F(4, 252) = 27.35, p < .001, \eta^2_p = .30$, confirming that the effects of increasing sentence lengths on appropriateness ratings depend on each length’s relative ranked position. This interaction is graphed in Fig. 2; the 95% confidence intervals for a group that do not bound the mean of the other group indicate statistically significant difference—hence, as expected, participants’ ratings of the 2nd and 4th common points differed, whereas the ratings of the 1st, 3rd and 5th common points did not. All the other effects, including the 3-way interaction, were not significant (all $F_s < 1$) and confirmed that rank effects were observed for both crime scenarios.

The observed context effects supported a relative rank account of sentencing evaluations—as sentences available in the immediate context biased participants’ assessments. As predicted, the same sentence was judged differently depending on its rank position within the sentence lengths set, supporting the rank principle. Other holistic features of the context (e.g., average length) did not play a role, thus no support was observed for alternative theories (e.g., ALT; Helson, 1947).

Table 3

Sentences lengths (in months) used for Study 2. Underlined lengths represent the five common points for the comparison between unimodal and bimodal distributions.

Crime scenario (and severity)			
Fraud (Low)		Sexual offenses (High)	
Distribution type			
Unimodal	Bimodal	Unimodal	Bimodal
3	3	12	12
8	4	30	15
9	5	34	19
10	6	38	23
11	8	41	30
12	12	45	45
13	16	49	60
14	18	53	68
15	19	55	71
16	20	60	74
21	21	78	78

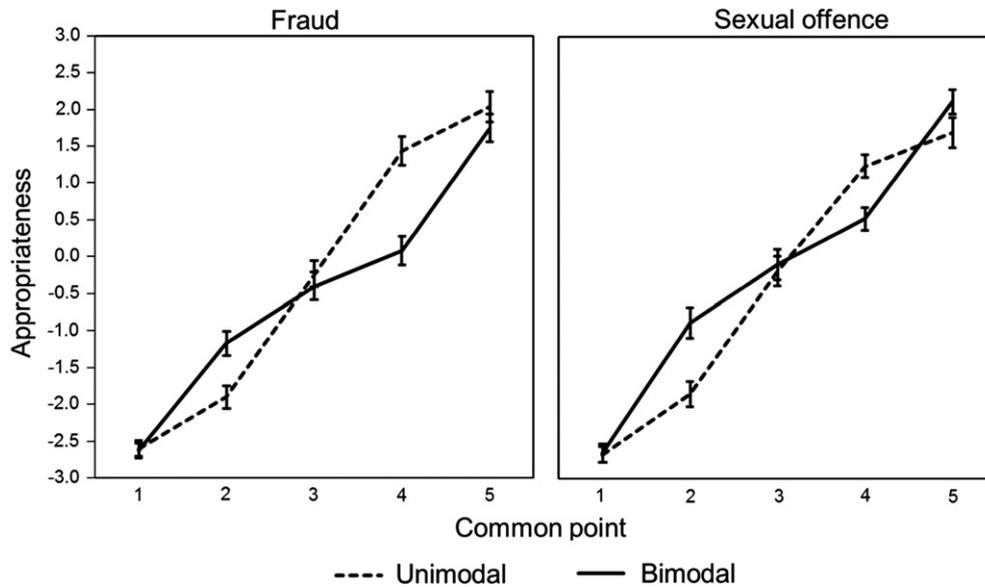


Fig. 2. Interactions between distribution and common points (1 to 5) for the two crime scenarios (Study 2). Error bars represent standard error of the mean.

4. Study 3

In this study, we experimentally tested the influence of the range principle on sentencing evaluations. A corollary of the range principle is that contextual skew, the clustering of stimuli at either end of the distribution, will affect the mean evaluation of the stimuli even if the actual mean of the stimuli is held constant. Negatively skewed distributions are usually rated higher on the dimension under consideration than positively skewed distributions, even when the distribution means are the same (e.g., Parducci, 1995). This results from the fact that in the negatively skewed distribution most stimuli cluster near the top end of the distribution and hence will receive relatively high evaluations according to the range principle. On the other hand, stimuli in the positively skewed distributions are mostly near to the lower limit of the distribution, and hence will attract lower subjective evaluations.

We experimentally manipulated the distribution skew of 11 sentence lengths and asked participants to indicate whether they would consider each of a given set of 11 sentences as too lenient, or not. Although the average sentence length was the same for both distributions (i.e. 12 months for fraud and 45 months for sexual offense, as in Study 2), we expected participants who were presented with a negatively skewed distribution to accept longer sentences as appropriate.

4.1. Method

4.1.1. Participants

A total of 79 participants (57 females) were recruited at a university in the UK to take part in this study. Those participants who were students were rewarded with course credits in exchange for their participation. Participants' ages ranged from 18 to 68 years ($M = 23.31$, $SD = 11.37$). Participants were predominantly of White ethnicity (59.49%), while a minority was of Chinese ethnicity (15.19%).

4.1.2. Design and procedure

Participants filled in a 2-page booklet which described fraud and sexual offense crimes, as in Study 2. The participants' task was to consider 11 sentence lengths (whose distribution was either positively or negatively skewed; see Fig. 3) to indicate their opinion on whether the Attorney General of England and Wales would have judged each sentence length as either too lenient or appropriate.

We first introduced the role of the Attorney General of England and Wales, who has the power to ask the Court of Appeal to review given sentences. Participants were told that "the Attorney General must consider the case personally and decide whether a particular sentence is unduly lenient in accordance with section 36(1)(a). That condition may be satisfied if the judge has erred in law as to his/her powers of sentencing (section 36(2)) to the extent that the resulting sentence is unduly lenient. The Attorney General must exercise his/her discretion in respect of the individual case. If the Court of Appeal agrees with the Attorney General, it can increase the sentence."

Participants were then presented with 11 sentence lengths. For each, participants had to indicate whether they thought that the Attorney General should refer them for being too lenient; participants had to tick either a box labeled "Yes, refer to Court of Appeal as too lenient" or a box labeled "No, do not refer to Court of Appeal" which were both displayed next to each sentence length.

The dependent variable in this study was the shortest sentence which was deemed by the participant to be appropriate (and hence not referred as too lenient). Some participants failed to comprehend the instructions as they referred relatively longer sentences as too lenient but did not refer shorter ones (10.13% for the low severity scenario, 11.39% for the high severity scenario); data from these participants were removed. Data removal did not compromise the balance of the design.

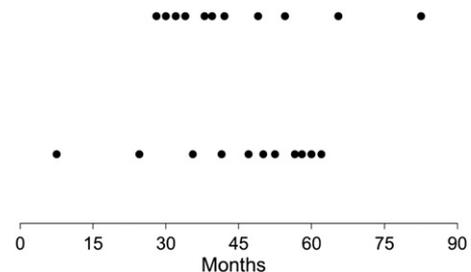


Fig. 3. Sentences lengths (in months) used for sexual offense scenario in Study 3. The top line of black dots represents the positively skewed distribution, while the bottom one is negatively skewed (most values cluster at the upper end of the distribution). The distributions for the fraud scenarios were the same, with the only difference being that they were centered at 12 instead of 45 months.

4.2. Results and brief discussion

Table 4 represents the average sentence length that was deemed appropriate for both the negatively and positively skewed conditions, separately for each crime scenarios. In line with the range principle, this sentence length was longer in the negatively skewed distribution, this finding holding true for both crime scenarios.

As the dependent variable consisted of one value for each scenario for each participant, non-parametric analyses were carried out. Two separate Mann–Whitney *U* tests were run and confirmed the observations above. For both the fraud crime scenario, $U(71) = 864.00$, $z = 2.70$, $p = .007$, $r = .32$, and sexual offense scenario, $U(70) = 814.00$, $z = 2.39$, $p = .017$, $r = .29$, the average appropriate sentence was higher in the negatively skewed distribution than in the positively skewed distribution.

As in Study 2, participants showed great sensitivity to the immediate context when deciding upon the appropriateness of sentence length. Participants' sentencing judgments were influenced by the range principle.

5. General discussion

The results of the present three studies largely support a rank-based account of sentencing evaluations, with additional contributions from range, and suggest that DbS (Stewart et al., 2006) can be applied successfully to a range of decision-making scenarios in the legal domain. Study 1 showed that people's subjective judgments of the seriousness of, and their concern about, their own law infringements depend largely on their beliefs about other people's dishonest behavior. Specifically, their judgments were determined by the relative ranked position of the frequency of their own illegal downloading copyrighted materials over the internet within the frequency distribution of downloading done by others. This result is consistent with social norms theory and with the research of Mazar, Ariely and their colleagues (Mazar & Ariely, 2006; Mazar et al., 2008); indeed, of the larger pool of approached participants, a very large proportion (86.54%) admitted downloading copyrighted materials from the internet—but to a small degree (the average number of downloaded items was just over 20 per month). However, here we provide a different theoretical perspective to account for these findings. In their research on cheating, the Mazar, Ariely and their colleagues showed that most people acted dishonestly, even if only to a small extent. The authors proposed a theory of self-concept maintenance, whereby people do internalize the norms and values of their society—as proposed by social norms theorists. Thus, when people consider the utility and disutility for being dishonest, they assess their behavior against an internalized standard which reflect their perception of society. If one complies with such moral system, one is rewarded—otherwise, one lowers her own self-image. When one person is torn between the gains of cheating and the threat to the self-concept as an honest person, she makes uses of two protective mechanisms which allows her to enjoy the first without compromising the second (e.g., inattention to own standards and malleability of the boundaries of the category of 'honest' behaviors; Mazar et al., 2008). Here instead we explain the findings of Study 1 by claiming

that participants' perception of the seriousness of their own law infringing depends on their (often inaccurate beliefs) about their own ranked position within the distribution of illegal downloaders in the UK; as participants inaccurately thought that about 84% of people do more illegal downloading than they do, the perceived seriousness of their own law infringements was relatively low, and the related attitudes rather permissive. More generally, people seem to hold inaccurate beliefs about crime rates in general (McIntyre, 1967; Roberts, 1992; Stalans & Lurigio, 1990; but see Marr, 1982)

A practical implication that can be suggested on the basis of the present findings concerns how the information should be presented in interventions aiming at reducing illegal downloading. Providing information about the average amount of downloading done in general might not be an efficient way to influence people's decision to download copyrighted material over the internet. As people compare themselves to others in terms of rank positions, rank-relevant information (e.g., percentiles) could prove to be more useful in deterring people from illegal downloading, as it is more in line with the way people process information about the frequency of dishonest behavior.

A complementary account could approach the direction of causality between beliefs, attitudes and dishonest behavior from a different angle. Participants could adjust the beliefs about other people's frequency of dishonest behavior so not to tarnish their own self-image as 'honest' people. As the seriousness of a crime committed by the respondents is determined by the relative ranked position of such crime among the inferred distribution of other people dishonest behavior, by shifting upwards the inferred distribution of downloading done in the UK, one decreases the perceived seriousness of her own crime (for similar arguments, see Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995; Allison, Messick, & Goethals, 1989; Schweitzer & Hsee, 2002). The suggestion here is that people might accommodate their beliefs about reality in order to 'justify' their own behavior (see also Bem, 1972; Festinger, 1957).

Relative comparisons are at odds with the assumption underlying some of the indexes of crime severity that have been compiled (e.g., Sellin & Wolfgang, 1964). It is sometimes assumed, explicitly or implicitly, that there is a direct relationship between the injury component of a criminal act and the perceived seriousness of the act. For instance, in the Sellin–Wolfgang Index of Crime Severity (1964), the seriousness score for the dimension of property theft is directly related to the value of the stolen property: A stolen value of \$10 is associated with a score of 1, a value between \$10 and \$250 is scored as a 2, and so on. This method can be labeled as an *absolute* approach, in that it assumes a direct relationship between severity of a crime and its subjective perceived seriousness, and therefore minimizes the influence of context effects. Similarly, some perspectives on the issue of law infringements and their perceived seriousness can be framed as absolute accounts. According to the standard economic approach, when people assess whether to engage in a dishonest act—e.g., whether or not to break the law—they perform the appropriate trade-off between (a) the anticipated reward as a consequence of the dishonest act, (b) the perceived likelihood of getting caught and (c) the expected punishment if they do indeed get caught. Thus, rates of dishonest behavior will increase as the expected benefits increase—while they decrease as the perceived likelihood of being punished, and the punishment itself, increase (e.g., Becker, 1968, 1976).

The information available in the decision-making context influenced participants' evaluations of sentencing. Rank and range principles were validated experimentally and were shown to drive judgments on the appropriateness of given sentences. This finding suggests that public consensus over crime seriousness and sentences appropriateness might be undermined not only by features of the respondents or other extra-evidential information (e.g., Stylianou, 2003), but also by cognitive processes which are vulnerable to context biases. In focusing exclusively on the role of comparison context, the present approach contrasts with and extends the work of von Helversen and Rieskamp (2009), who tested

Table 4
Mean participants' responses as a function of crime severity and distribution skew (Study 3).

Distribution	Negative skew		Positive skew		Effect size <i>r</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Crime					
Fraud	12.50	2.22	10.96	1.94	0.32
Sexual offense	48.12	9.05	44.74	9.06	0.29

the predictions of different cognitive models (e.g., disjunctive and conjunctive models; Einhorn, 1970; the mapping model; von Helversen & Rieskamp, 2008) of how people integrate various potentially aggravating or mitigating factors (e.g., number of previous convictions) to reach a sentencing decision.

One consequence is that people might not have stable internal standards, a sort of 'internal punishment' scale where the crime under consideration is mapped directly onto its correspondent punishment—after having made adjustments due to circumstantial factors. Here we argue that the boundaries on this internal map are rather large and variable, and that reference information in the decision-making context might alter these boundaries to a substantial degree. The fact that anchoring effects are observed in judgments about legal matters supports this conclusion (e.g., Englich & Mussweiler, 2001; Englich, Mussweiler, & Strack, 2005, 2006; Mussweiler & Englich, 2005; Robbennolt & Studebaker, 1999). Anchoring effects are observed when people's evaluations are influenced by a prior quantity (the 'anchor'), which could for example be represented by the lowest and highest values in the context (i.e. the range). In the courtroom, anchoring effects can lead to disparate sentences being given for identical crimes. For instance, sentencing demand by the prosecutors has a direct influence on the judge's sentence—which is assimilated to the sentence demanded by the prosecutor. Strikingly, these anchoring effects appear to be independent of (a) how relevant the prosecutor's demand was perceived to be by the judge and (b) the judge's experience (Englich et al., 2005).

The present model addresses the current debate on context effects in legal decision-making by exploring the nature of the cognitive comparisons behind such effects. Of course, the model cannot account for every contextual bias observed in legal decision-making. For instance, it does not directly address the question on how people integrate different pieces of information relevant to the case under consideration in order to make a sentencing evaluation (Dhami & Ayton, 2001; von Helversen & Rieskamp, 2009). Moreover, the rank-based model does not directly explain the assimilation effects central to the anchoring phenomenon (e.g., Englich et al., 2005) nor the contrast effects observed for people's evaluation of the severity of crimes (e.g., Pepitone & DiNubile, 1976). For the latter, people's evaluations of the severity of a crime depend partially on how they evaluated previously presented wrongdoings. For instance, a case of homicide was regarded as less serious if the participants had previously rated another homicide, compared to when the very same case of homicide was evaluated after a case of assault (for a comprehensive discussion on contrast effects in the law domain, see Rachlinski & Jourden, 2003). It is also important to discuss the factors that may limit the applicability of the present model. First, the model requires the presence of quantities in the decision-making context. Indeed, in all the three present studies, participants had to provide evaluations based on numerical values, either referring to sentences handed out for a crime or to the amount of items illegally downloaded over the internet. However, Stewart et al. (2006), acknowledging that people prefer verbal rather than numerical descriptions of probabilities, test and validate the assumptions of DbS also for non-numerical quantities (such as description of occurrence like 'very likely' and 'small chance'). Thus, future studies could investigate whether the rank-based model can explain context effect when words, rather than numbers, are used to express the occurrence of specific behaviors in the legal domain. Second, in Studies 2 and 3 the information was provided in tables—all the contextual information was present at the time participants had to make their evaluations. Future research could investigate contextual biases arising from sequential presentation, whereby the to-be-evaluated targets are provided one at a time.

In conclusion, this paper proposes a relative ranked account of people perception of crime seriousness and sentencing fairness. The model describes the cognitive processes underlying judgments and decision-making in the legal domain, illustrating how the same cognitive dynamics may underlie context effects in different judicial domains.

Acknowledgments

This work was supported by the Economic and Social Research Council [grant numbers RES-062-23-2462, ES/K002201/1] and the Leverhulme Trust [grant number RP2012-V-022].

References

- Alicke, M.D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, 68, 804–825.
- Allison, S. T., Messick, D.M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition*, 7, 275–295.
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76, 169–217.
- Becker, G. S. (1976). *The economic approach to human behavior*. Chicago, IL: University of Chicago Press.
- Bem, D. J. (1972). Self perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 1–62). New York: Academic Press.
- Boyce, C. J., Brown, G. D. A., & Moore, S.C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological Science*, 21, 471–475.
- Brown, G. D. A., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does wage rank affect employees' well-being? *Industrial Relations*, 47, 355–389.
- Bushway, S. D., Owens, E. G., & Piehl, A.M. (2012). Sentencing guidelines and judicial discretion: Quasi-experimental evidence from human calculation errors. *Journal of Empirical Legal Studies*, 9, 291–319.
- Campbell, E. Q. (1964). The internalization of moral norms. *Sociometry*, 27, 391–412.
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58, 1015–1026.
- Dhami, M. K. (2003). Psychological models of professional decision making. *Psychological Science*, 14, 175–180.
- Dhami, M. K., & Ayton, P. (2001). Bailing and jailing the fast and frugal way. *Journal of Behavioral Decision Making*, 14, 141–168.
- Durham, A.M. (1993). Public opinion regarding sentences for crime: Does it exist? *Journal of Criminal Justice*, 21, 1–11.
- Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, 73, 221–230.
- Eisenstein, J., & Jacob, H. (1977). *Felony justice: An organizational analysis of felony courts*. Boston, MA: Little Brown and Company.
- Englich, B., & Mussweiler, T. (2001). Sentencing under uncertainty: Anchoring effects in the courtroom. *Journal of Applied Social Psychology*, 31, 1535–1551.
- Englich, B., Mussweiler, T., & Strack, F. (2005). The last word in court—A hidden disadvantage for the defense. *Law and Human Behavior*, 29, 705–722.
- Englich, B., Mussweiler, T., & Strack, F. (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, 32, 188–200.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Gebotys, R. J., Roberts, J. V., & DasGupta, B. (1988). News media use and public perception of crime seriousness. *Canadian Journal of Criminology*, 30, 3–16.
- Gerbas, K. C., Zuckerman, M., & Reis, H. T. (1977). Justice needs a new blindfold: A review of mock jury research. *Psychological Bulletin*, 84, 323–345.
- Graber, D. (1980). *Crime news and the public*. New York, NY: Praeger.
- Hagerty, M. R. (2000). Social comparisons of income in one's community: Evidence from national surveys of income and happiness. *Journal of Personality and Social Psychology*, 78, 764–771.
- Hamilton, V. L., & Rytina, S. (1980). Social consensus on norms of justice: Should the punishment fit the crime? *The American Journal of Sociology*, 85, 1117–1144.
- Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *The American Journal of Psychology*, 60, 1–29.
- Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychological Review*, 55, 297–313.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C. F., Fehr, E., & Gintis, H. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *American Economic Review*, 91, 73–78.
- Hoffman, P. B., & Hardyman, P. L. (1986). Crime seriousness scales: Public perception and feedback to criminal justice policymakers. *Journal of Criminal Justice*, 14, 413–431.
- Janiszewski, C., & Lichtenstein, D. R. (1999). A range theory account of price perception. *Journal of Consumer Research*, 25, 353–368.
- Johnson, B.D. (2006). The multilevel context of criminal sentencing: Integrating judge- and county-level influences. *Criminology*, 44, 259–298.
- Kwan, Y. K., Chiu, L. L., Ip, W. C., & Kwan, P. (2002). Perceived crime seriousness. Consensus and disparity. *Journal of Criminal Justice*, 30, 623–632.
- Lewandowsky, S., Griffiths, T. L., & Kalish, M. L. (2009). The wisdom of individuals: Exploring people's knowledge about everyday events using iterated learning. *Cognitive Science*, 33, 969–998.
- Maltby, J., Wood, A.M., Vlaev, I., Taylor, M. G., & Brown, G. D. A. (2012). Contextual effects on the perceived health benefits of exercise: The exercise rank hypothesis. *Journal of Sport & Exercise Psychology*, 34, 828–841.
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72, 1329–1376.
- Marr, M. (1982). The accuracy of public belief about crime. *Criminology*, 20, 185–204.
- Mazar, N., & Ariely, D. (2006). Dishonesty in everyday life and its policy implications. *Journal of Public Policy and Marketing*, 25, 117–126.

- Mazar, N., On, A., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45, 633–644.
- McIntyre, J. (1967). Public attitudes toward crime and law enforcement. *The Annals of the American Academy of Political and Social Science*, 374, 34–46.
- Mellers, B.A. (1982). Equity judgment: A revision of Aristotelian views. *Journal of Experimental Psychology. General*, 111, 242–270.
- Mellers, B.A. (1986). Fair allocation of salaries and taxes. *Journal of Experimental Psychology. Human Perception and Performance*, 12, 80–91.
- Melrose, K. L., Brown, G. D. A., & Wood, A.M. (2013). Am I abnormal? Relative rank and social norm effects in judgments of anxiety and depression symptoms severity. *Journal of Behavioral Decision Making*, 26, 174–184.
- Ministry of Justice (2010). *Sentencing statistics. England and Wales 2008. Statistical Bulletin*. London: M. o. Justice.
- Mussweiler, T., & Englich, B. (2005). Subliminal anchoring: Judgmental consequences and underlying mechanisms. *Organizational Behavior and Human Decision Processes*, 98, 133–143.
- Niedrich, R. W., Weathers, D., Hill, C. R., & Bell, D. R. (2009). Specifying price judgments with range-frequency theory in models of brand choice. *Journal of Marketing Research*, 46, 693–702.
- O'Connell, M., & Whelan, A. (1996). Taking wrongs seriously. Public perceptions of crime seriousness. *British Journal of Criminology*, 36, 299–318.
- Orviska, M., & Hudson, J. (2002). Tax evasion, civic duty, and the law abiding citizen. *European Journal of Political Economy*, 19, 83–102.
- Parducci, A. (1965). Category judgment: A range-frequency model. *Psychological Review*, 72, 407–418.
- Parducci, A. (1995). *Happiness, pleasure, and judgment: the contextual theory and its applications*. Mahwah, NJ: L. Erlbaum Associates.
- Parducci, A., Calfee, R. C., Marshall, L. M., & Davidson, L. P. (1960). Context effects in judgment: Adaptation level as a function of the mean, midpoint, and median of the stimuli. *Journal of Experimental Psychology*, 60, 65–77.
- Parducci, A., & Perrett, L. F. (1971). Category rating scales: Effects of relative spacing and frequency of stimulus values. *Journal of Experimental Psychology*, 89, 427–452.
- Peer, E., & Gamliel, E. (2013). Heuristics and biases in judicial decisions. *Court Review* (in press).
- Pepitone, A., & DiNubile, M. (1976). Contrast effects in judgments of crime severity and the punishment of criminal violators. *Journal of Personality and Social Psychology*, 33, 448–459.
- Posner, E. A. (2000). Law and social norms: The case of tax compliance. *Virginia Law Review*, 86, 1781–1819.
- Rachlinski, J. J., & Jourden, F. (2003). The cognitive components of punishment. *Cornell Law Review*, 88, 457–485.
- Robbenolt, J. K., & Studebaker, C. A. (1999). Anchoring in the courtroom: The effects of caps on punitive damages. *Law and Human Behavior*, 23, 353–373.
- Roberts, J. V. (1992). Public opinion, crime, and criminal justice. *Crime and Justice*, 16, 99–180.
- Roberts, J. V. (2003). Public opinion and mandatory sentencing: A review of international findings. *Criminal Justice and Behavior*, 30, 483–508.
- Roberts, J. V., & Doob, A. N. (1989). Sentencing and public opinion: Taking false shadows for true substances. *Osgoode Hall Law Journal*, 27, 491–515.
- Roberts, J. V., & Stalans, L. (1997). *Public opinion, crime, and criminal justice*. Boulder, CO: Westview Press.
- Rossi, P. H., Simpson, J. E., & Miller, J. L. (1985). Beyond crime seriousness: Fitting the punishment to the crime. *Journal of Quantitative Criminology*, 1, 59–90.
- Ruback, R. B., & Wroblewski, J. (2001). The federal sentencing guidelines: Psychological and policy reasons for simplification. *Psychology, Public Policy, and Law*, 7, 739–775.
- Schweitzer, M. E., & Hsee, C. K. (2002). Stretching the truth: Elastic justification and motivated communication of uncertain information. *Journal of Risk and Uncertainty*, 25, 185–201.
- Sellin, T., & Wolfgang, M. E. (1964). *The measurement of delinquency*. Montclair: Patterson Smith.
- Smith, R. H., Diener, E., & Wedell, D. H. (1989). Intrapersonal and social-comparison determinants of happiness—A range-frequency analysis. *Journal of Personality and Social Psychology*, 56, 317–325.
- Stalans, L. J. (1993). Citizens' crime stereotypes, biased recall and punishment preferences in abstract cases: The educative role of interpersonal sources. *Law and Human Behavior*, 17, 451–470.
- Stalans, L. J., & Diamond, S. S. (1990). Formation and change in lay of criminal sentencing evaluations: Misperceptions and discontent. *Law and Human Behavior*, 14, 199–214.
- Stalans, L. J., & Lurigio, A. J. (1990). Lay and professionals' beliefs about crime and criminal sentencing: A need for theory, perhaps schema theory. *Criminal Justice and Behavior*, 17, 333–349.
- Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, 112, 881–911.
- Stewart, N., Chater, N., & Brown, G. D. A. (2006). Decision by sampling. *Cognitive Psychology*, 53, 1–26.
- Stylianou, S. (2003). Measuring crime seriousness perceptions: What we learned and what else do we want to know. *Journal of Criminal Justice*, 31, 37–56.
- Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy*, 26, 89–103.
- Volkman, J. (1951). Scales of judgment and their implications for social psychology. In J. H. Rohrer, & M. Sherif (Eds.), *Social psychology at the crossroads*. : New York Harper & Row.
- von Helversen, B., & Rieskamp, J. (2008). The mapping model: A cognitive theory of quantitative estimation. *Journal of Experimental Psychology: General*, 137, 73–96.
- von Helversen, B., & Rieskamp, J. (2009). Predicting sentencing for low-level crimes: Comparing models of human judgment. *Journal of Experimental Psychology: Applied*, 15, 375–395.
- Warr, M. (1989). What is the perceived seriousness of crime? *Criminology: An Interdisciplinary Journal*, 27, 795–821.
- Watkinson, P., Wood, A.M., Lloyd, D., & Brown, G. D. A. (2013). Pain ratings reflect cognitive context: A range frequency model of pain. *Pain*, 154, 743–749.
- Wood, A.M., Boyce, C. J., Moore, S.C., & Brown, G. D. A. (2012a). An evolutionary based social rank explanation of why low income predicts mental distress: A 17 year cohort study of 30,000 people. *Journal of Affective Disorders*, 136, 882–888.
- Wood, A.M., Brown, G. D. A., & Maltby, J. (2011). Thanks, but I'm used to better: A relative rank model of gratitude. *Emotion*, 11, 175–180.
- Wood, A.M., Brown, G. D. A., & Maltby, J. (2012b). Social norm influences on evaluations of the risks associated with alcohol consumption: Applying the rank based Decision by Sampling model to health judgments. *Alcohol and Alcoholism*, 47, 57–62.
- Wood, A.M., Brown, G. D. A., Maltby, J., & Watkinson, P. (2012c). How are personality judgments made? A cognitive model of reference group effects, personality scale responses, and behavioral reactions. *Journal of Personality*, 80, 1275–1311.
- Yankelovich, S. (1984). *Taxpayer attitudes study—Final report* (Washington, DC).