



# The perceived seriousness of crimes: inter-individual commonalities and differences

Ingwer Borg<sup>1</sup> · Dieter Hermann<sup>2</sup> · Wolfgang Bilsky<sup>1</sup>

Accepted: 16 March 2022 / Published online: 6 April 2022  
© The Author(s) 2022

## Abstract

The paper focuses on ratings of the perceived seriousness of crimes and how they can be explained in a testable scaling model (SOCID) that shows commonalities and differences among individuals. The model assumes that crimes are first represented on each person's psychological map. This map can be described as a geometric space spanned by the crimes' perceived attributes. When assessing the seriousness of crimes, the person positions him-/herself on this map by finding a point whose distances to the crime points express how much he/she is rejects the respective crimes. Using three representative German surveys on crime prevention, the SOCID model is positively tested for various predictions: (1) All persons in our surveys share the same psychological map of crimes; (2) a 2-dimensional map accurately describes the persons' observed badness ratings; (3) the points representing the crimes form essentially linear manifolds on the maps; (4) the older and the more conservation-oriented a person, the more similar his/her distances to the crimes; (5) gender has no systematic effect on the distribution of person-points on the crimes' maps.

**Keywords** Seriousness of crimes · Norm acceptance · Unfolding · Inter-individual differences · Age · Personal values

## 1 Introduction

People's perception of the seriousness of crimes has been a topic of great interest in the social sciences. How people feel about crimes is typically assessed by surveys that ask the respondents to rate the "badness" or "wrongness" of various offenses such as smoking pot, tax evasion, or robbery on some answer scale (Stylianou 2003; Adriaenssen et al. 2017). The scales used for such ratings are, for example, a 0-to-10 rating scale with labels ranging from "not bad at all" (0) to "very bad" (10); a scale with the categories "bad", "don't care", "good"; or a permissiveness scale with categories "would/would not tolerate". The

---

✉ Ingwer Borg  
ingwer.borg@gmail.com

<sup>1</sup> Fachrichtung Psychologie, Westfälische Wilhelms-Universität, Münster, Germany, Fliednerstr. 21, NRW 46149

<sup>2</sup> Institut für Kriminologie, Ruprecht-Karls-Universität, Heidelberg, Germany, Friedrich-Ebert-Anlage 6-10, BW 69117

persons' numerically coded ratings on such scales are often interpreted as indicators of their norm acceptance vis-à-vis the particular offense, and averaging the ratings of a person across items is taken as a person's general acceptance of legal norms and as a predictor of sanctions-based behavioral expectations (Dietrich 2017; Seddig 2014, 2016; Borg and Herrmann 2021).

The classic study of peoples' perception of the seriousness of crimes is an investigation by Sellin and Wolfgang (1964). They demonstrated remarkable consensus across social groups and countries in how people assess 141 crimes -- described in contextual detail such as: "The offender stabs a victim with a knife [and] the victim is treated by a physician but requires no further treatment." -- in terms of their seriousness. This landmark study led to a series of follow-up investigations that focused on various substantive and methodological issues (see Stylianou (2003), for a comprehensive review). We will look at some of these issues that are particularly important for understanding the psychology of an individual's judgments on the seriousness of crimes.

### 1.1 Two issues of seriousness-of-crimes measurements

One important issue is what characteristics of crimes affect their perceived seriousness. This question was addressed in different ways in the literature. One attempt was first classifying crimes by facets such as their degree of violence, the amount of harm they cause, the nature of harm (economic, physical), the type of victim, etc., and then using regression analysis to explain the observed seriousness ratings by these facets. It was found, for example, that violent crimes are rated as more serious than property crimes, and that crimes with more severe consequences receive higher seriousness ratings than those with minor consequences (Sellin and Wolfgang 1964; Rossi et al. 1991; Warr 1989).

Another approach was asking the respondents to not only rate crimes on their seriousness but also on such criteria as immorality, wrongfulness, or the severity of harm caused by the crime, and then statistically explaining the seriousness ratings by these external variables. Yet, with items that elaborate the crimes' scenarios in detail, one can ask to what extent this *evokes* or *induces* perceptual dimensions that influence the respondents' seriousness judgments. Such item-induced features may not be relevant in everyday settings. This becomes even more likely if the respondents are explicitly asked to rate the crimes not only on their seriousness but also on their wrongfulness or on the harm they cause. The effect that such additional scales may have on the seriousness ratings is potentially further enhanced by collecting ratings on external scales *before* the seriousness ratings, and by even asking the respondents to *explicitly consider* such criteria in their seriousness judgments: "Respondents were instructed to reflect on the components of seriousness by asking: 'Taking into account the wrongfulness of the crime, the severity of its harms and the incidence of the crime and its harms, how serious do you think the following crimes are?' " (Adriaenssen et al. 2017, p. 135).

It could be also be argued that elaborating crime scenarios along certain dimensions or using additional scalings of crimes on explanatory criteria impacts serious-of-crimes judgments in the sense that they become more *analytical* rather than *integral*. In a context, where fast and frugal answers without any consequences are collected (as in anonymous population surveys), the respondents are likely to use *simple heuristics* to generate their responses (Gigerenzer and Gaissmaier 2011). Moreover, first suggesting a set of criteria for seriousness judgments, may have an effect on the variance across individuals, leading to more consensus because all persons are primed to use the same or similar criteria.

Consensus is generally a major topic of research on the perception of the seriousness of crimes. One can distinguish consensus at the level of the individuals as units of analysis or the level of groups (e.g., men vs. women, age groups, countries, historical groups, socio-economic groups, religion, white collar vs. blue collar, minorities). Stylianou (2003) gives an overview of studies comparing groups (varying on gender, race, country, education, income, etc.) in terms of their seriousness ratings. In summary, these studies show much overall agreement of the seriousness scales across various demographic groups in representative surveys, but, of course, these studies are not exhaustive in terms of individual differences. Many additional psychological criteria could be used to study inter-group consensus. Indeed, "there can always be additional factors affecting individual judgments" (Stylianou 2003, p. 40).

Differences among subgroups are also reported but they are limited to *special* items and particular subgroups such as tax evasion or rape as assessed by men and by women. Such items take on somewhat different positions on the general seriousness scales for these groups, but the overall scales remain quite similar. Yet, Sellin and Wolfgang (1985, p. 7) remark that "the strong agreement among subgroups might effectively obscure any individual differences".

There are few studies on that question. One example is Rossi et al. (1991). These authors correlated each individual's ratings with the mean ratings of the entire sample for each crime: The mean correlation was .54. Thus, there is agreement but only in this simple item-by-item overall correlational sense. Inter-individual differences are left to unexplained variance. There are no studies that explain inter-individual differences in a comprehensive model. Moreover, even if such studies converge to a robust overall seriousness scale that is shared by many, most, or all individuals, it remains unclear whether it would correspond to the scales usually reported in the literature.

It is also unclear to what extent the seriousness ratings do indeed form *uni- or multi-dimensional* scales. The scalability of the ratings is rarely tested even though the crimes used in previous studies can be classified by many different criteria, dimensions, design factors, or facets. This leads to many crime types with profiles that are not always comparable among each other. Hence, it is at least possible that a particular set of crimes is not uni-dimensionally scalable. Indeed, Abrams and Della-Fave (1976) report that seriousness ratings on victimless crimes (such as prostitution, drug use, gambling, etc.) led to *multiple* factors when factor analyzed.

In the following, we present a model that seeks answers to the above questions at the level of the individual. The (testable) model explains how an individual arrives at his/her seriousness judgments based on a mental representation of crimes that is shared by all individuals in a given sample.

## 1.2 SOCID: an unfolding model of seriousness-of-crimes judgments

The basic assumption of our model is that we interpret a person's seriousness judgment as a distance function. A person who rates a certain crime as "very serious", for example, clearly *distances* him- or herself from the crime as he/she perceives it. Crimes are represented as points in the person's psychological space -- the crimes' mental representation -- spanned by the crimes' attributes. Such attributes can be facets, dimensions, class criteria, or characteristics as those described by Sellin and Wolfgang (1964), Rossi et al. (1991) or McCleary et al. (1981). That is, for example, the amount of violence exerted by the crime, the type of victim (single person, group, small/large company, an organization,

the State), or the extent to which the crime is committed willfully and deliberately. Every person activates or constructs his/her own perceptual crime space when asked to assess the seriousness of crimes, and then “measures” how much he/she distances him-/herself from the crime points in this crime space. These measurements are then “edited” to fit into the format of the response scale provided by the researcher (Tourangeau et al. 2003; Schwarz 2008).

More formally, consider the observed person-by-seriousness rating matrix. We denote this matrix as  $\Delta$ , with elements  $\delta_{ij}$ .  $\Delta$  has the order  $N \times n$ , with  $N$  persons and  $n$  crimes. We assume that  $\delta_{ij}$  corresponds to person’s  $i$  psychological distance from crime  $j$ . That is, for example, given a uni-polar 10-point seriousness rating scale, a person who scores 8, say, for crime  $X$ , expresses by this score that she clearly distances herself from  $X$ . If, on the other hand, a person scores 2, say, it means that he does not dissociate himself very much from  $X$ . This corresponds closely to the “unfolding” model of Coombs (1964), except that we interpret the observed  $\delta$ ’s as ratio-scaled distance estimates and not as ordinal indicators of a persons’ preferences. For short, we call this model SOCID for S(eriousness)O(f)C(rimes)I(individual)D(ifferences) model.

The notion “individual differences model” is a standard concept in the scaling literature (Borg and Staufenbiel 2007; Dunn-Rankin et al. 2004). Two well-known examples of such models are the INDSCAL model and its variants (Carroll and Chang 1970; Borg and Groenen 2005; Borg et al. 2018) and the unfolding model (Coombs 1964; Busing et al. 2010). Individual differences models represent not just the stimuli or items but also every single person, person by person. That is, these models scale all data and not just an aspect of the data, such as the inter-correlations among the columns of the persons-by-items data matrix. If such models hold, they show a structure of the items that is shared by all persons. For example, in the original INDSCAL model, all persons perceive the items (or: stimuli) in the same way as a configuration of points in a fixed set of dimensions. However, they attribute different weights to these dimensions, thereby squeezing or stretching the psychological map in the directions of the common dimensions. Typically, such models also serve as psychological models that explain how an individual arrives at his/her judgments. Hence, they are not just statistical models that “describe” or “visualize” the observations. Indeed, such models are far from being descriptive only: They are testable and, if they hold, uncover an underlying psychological lawfulness.

Individual differences are often not of direct interest in individual differences models, although one may want to relate them to demographic variables such as age or psychological variables such as Big5 measurements or personal values. What is often the major issue of interest is showing whether a structure that has been found by, say, analyzing correlations *across* individuals also holds *within* individuals (Borg et al. 2017).

The SOCID model proposes that all individuals share a common seriousness-of-crimes scale but that they differ in how they position themselves, person by person, with respect to this scale. We show that this model holds and that it is surprisingly simple. Age, gender, and personal values can be used to check the substantive validity of the model because these variables have been found—in aggregate-based research—to be moderators of seriousness-of-crimes judgments.

We aim to optimally represent each  $\delta_{ij}$  by the distance  $d_{ij}$  between a point for person  $i$  and a point for crime  $j$  in a joint  $m$ -dimensional Euclidean space. That is, we want to minimize

$$\sigma = \sum_{i=1}^N \sum_{j=1}^n (\delta_{ij} - d_{ij})^2, \quad (1)$$

where the distance between person-point  $i$  and crime point  $j$  is computed as

$$d_{ij} = \sqrt{\sum_{a=1}^m (x_{ia} - x_{ja})^2}, \quad (2)$$

with  $x_{ia}$  the coordinate of the point representing person  $i$  on dimension  $a$ , and  $x_{ja}$  the coordinate of the point representing crime  $j$  on dimension  $a$ . So, for example, if we have ratings of 1000 persons and 14 crimes, we want to optimally represent 14,000  $\delta_{ij}$  ratings by 14,000  $d_{ij}$  distances between 1000 person-points and 14 crime points in  $m$ -dimensional geometric space. This means that each single rating is represented in the model space without any pre-processing of the data (e.g., centering or normalizing), data aggregation, or data reduction. Every observed person is represented by one point, while there are only 14 points for the crimes. Note also that we do not allow for any “optimizing” transformations of the observations (except for an overall cosmetic scaling constant on the distances). Rather, we take the data seriously as distance estimates with a fixed origin that can be understood as “zero” (e.g. by using a label such as “not bad at all”). Thus we assume ratio-scale level of the rating data<sup>1</sup>.

Often, the structure of the points representing the objects in psychological space can be predicted based on substantive considerations. In case of crimes and seriousness ratings, such predictions depend on the particular set of crimes in the item sample. For sets comprising few crimes described in relatively abstract terms (such as “burglary” or “theft”), one plausible prediction is that they lead to a simple array that corresponds to the order of the crimes in terms of wrongfulness or harmful consequences (Stylianou 2003; Adriaenssen et al. 2017). Such a linear scale could be embedded in  $m$ -dimensional space and then each person could be fitted, one after the other, into this space by identifying points whose distances to the object points optimally approximate the observed seriousness ratings.

In contrast to this confirmatory approach, one could also *not* impose any theoretical restrictions onto the distribution of crime points and person-points. This exploratory approach allows the data to speak for themselves. In either case, the resulting configuration (in  $m$  dimensions) generates distances that may or may not explain the data, even if  $m$  is very large. Thus, the model is testable.

To measure the model fit, we compute the loss of the fitted model by

$$Stress = \sqrt{\sigma / \sum d_{ij}^2}. \quad (3)$$

Stress resembles the Stress-1 coefficient used in multidimensional scaling (Borg and Groenen 2005). It varies between 0 and 1. Perfect solutions have a Stress of zero. Non-perfect solutions have Stress values greater than 0. What one wants is a Stress value that is

<sup>1</sup> Weaker models that allow for interval or ordinal transformations of the data are easy to implement, but here we start with the strongest model. This ratio-scale model is directly interpretable and formally most testable.

acceptably small, statistically significant, and associated with a substantively interpretable solution.

### 1.3 Moderator variables in the SOCID model

External moderator variables of the respondents' seriousness ratings can also be studied in the SOCID model. They allow a to better understand the psychology of the model. Given that all persons share the same configuration of the crime points in psychological space, persons belonging to different groups may lead to different distributions of the person-points. Older persons, for example, tend to rate *all* crimes as relatively serious compared to young persons who are comparatively lenient towards petty offenses (Rossi et al. 1991; Borg 2021). Expressed in terms of the seriousness model, we would thus expect that older persons distance themselves more from *all* crimes than younger persons. Younger persons should be closer to crime points that represent petty offenses.

The effects of gender could be expressed similarly in the model. However, since there is no general evidence that women rate crimes generally higher than men (see Stylianou 2003), the SOCID model should show that one cannot reliably discriminate the person-points of men and women.

The SOCID model always assumes a common configuration of the crime points in model space for all individuals. If, for example, males rate violent crimes as more serious (relative to the other crimes) than women, and assault as relatively less serious (Walker 1978; McCleary et al. 1991), then this would require different model solutions in principle. However, as long as such *local* scale differences are not large, one common seriousness scale for all subgroups should be sufficiently precise.

Besides the standard demographic variables age and gender, an interesting psychological moderator of seriousness-of-crimes ratings is personal values. Personal values are broad goals that guide people's behavior in general directions. They have been studied extensively, mostly based on surveys asking people to rate how important it is for them to observe and strive for tradition, power, benevolence, and other such goals (see e.g. Schwartz 1992). Numerous studies have identified a set of ten "basic" values that seem universally valid for persons of any age, gender, and social background. These basic values form two sets of "higher-order" values that are opposed to each other on the value circle: Self-enhancement (composed of power and achievement values) vs. self-transcendence (benevolence, universalism), and conservation (security, conformity, tradition) vs. openness to change (hedonism, stimulation, self-direction). These higher-order values exhibit a clear relationship to crime ratings: Persons striving for conservation show generally higher badness-of-crimes ratings than persons who put more emphasis on openness to change; on the other hand, people's preferences for self-transcendence or self-enhancement are essentially unrelated to seriousness-of-crimes ratings (Bilsky et al. 2020; Borg and Hermann 2020). Thus, we predict that, in particular, persons who strongly strive for conservation will position themselves relatively far away from the crime points on the psychological map, because that makes the distances to the crime points relatively similar and relatively large.

## 2 Methods

### 2.1 Samples

All data come from a series of mail surveys on crime prevention in various German cities (for details, see Hermann 2017a, b; Hermann and Wachter 2017). We here use three samples that assess both the seriousness of crimes and also personal values with the same scales. The samples were representative random samples of city residents with a minimum age of 14 years, all drawn from the resident registers of the respective cities. All surveys were run as anonymous mail surveys. No incentives were offered to the respondents. Rather, they were asked in a letter signed by the mayor of the city to support the city administration by providing information needed to prevent crimes. The time window of the data collection was three weeks in all surveys. A thank-you note that also served as a reminder note (“in case you have not yet ...”) was also sent by mail to each potential respondent in the middle of the survey administration window.

The PF20 survey was conducted in 2019 on 8,000 persons in the city of Pforzheim. Its return rate was 28% ( $N=2,230$ ). Age was assessed by a category scale ranging from 14–19 years, 20–29 years, etc. to 80+ years. The mean age was 46 years.

The MA16 survey was conducted in 2016 on a total sample of 9,998 persons in the city of Mannheim. The participation rate was 36% ( $N=3,272$ ). Age was assessed using the 8-point scale from above. The mean age was 42 years.

The HD17 survey was run in 2017 in the city of Heidelberg on a total sample of 8,000 residents of the city. The response rate was 35% ( $N=2,770$ ). Age was assessed with the 8-point scale used in the other surveys. The average age was 45 years.

Return rates of about 30% are relatively high as this is a value that corresponds to the participation rate of the interviewer-based (!) European Social Survey in Germany (Jowell et al. 2007; Beullens et al. 2014). Moreover, all three realized samples closely match the population statistics in relevant demographic variables. Minor exceptions are that females are slightly over-represented in the surveys by about 5%, and older persons (aged 40 years or older) are over-represented by about 6%.

What should be remarked is that in each population and sample, the persons with a migration background was substantial. For example, the MA16 population comprised 35% individuals who were not born in Germany, but the realized sample showed only a proportion of 18%. In the other samples, citizenship was not assessed, but the respondents were asked about their “migrant background”, and 28% of the HD17 and 40% of the PF20 participants reported that they were either born outside of Germany or at least one of their parents was born abroad.

### 2.2 Instruments

The seriousness of crimes was measured with items that focused on offenses similar to those used in the ALLBUS 1990 (Allerbeck et al. 2017)<sup>2</sup>. The ALLBUS items were adapted for surveys conducted in the context of community crime prevention by Hermann

---

<sup>2</sup> The ALLBUS is a bi-annual representative survey monitoring trends in social attitudes and behavior in Germany. It regularly covers a wide spectrum of fixed standard topics and a set of special topics that change from survey to survey. The 1990 survey had a particular focus on deviant behavior and sanctions.

**Table 1** Offenses assessed for their seriousness; item 14 utilized in sample PF20 only

Nr.	Offense	Item
1	Fare evasion	Not paying the fare when using public transportation, fare dodging
2	Shoplift	Taking goods worth about 50 Euros in a store without paying
3	Drunk driving	Driving a car under the influence of more alcohol than allowed
4	Robbery	Ripping off someone's handbag
5	Cocaine	Taking cocaine
6	Smoke pot	Taking hashish
7	Assault	Beating or hitting someone without being in a self-defense situation
8	Vandalism	Smashing the windows of a phone booth
9	Tax fraud	Evading taxes if one has the opportunity to do so
10	Benefits fraud	Claiming sickness benefits, unemployment benefits or other social benefits even though one is not entitled to them
11	Car break	Breaking into someone's car to take it for a spin
12	Take bribes	Taking bribes
13	Hit and run	Not reporting the damage one has done to a parking car
14	Indecent touch	Indecent touching of another person

(2003). These items (see Table 1; item #14 used only in the PF20 sample, because of its relevance stirred by the #MeToo movement) focus on different offenses that vary in type and severity of norm violation. The scale does not contain items on major crimes (e.g., murder, rape, and arson), because such crimes were expected not to lead to sufficient variance in surveys using rating scales (see Wasmer et al. 1991; Hirtenlehner and Reinecke 2018; Bentrup et al. 2017).

The item battery was introduced by the following preamble: "Various forms of behaviors can be assessed differently. Please indicate whether you consider the following actions bad behaviors or not. 1 would mean that you consider the behavior not bad at all, and 7 that you consider it very bad".

Personal values were measured by the German version of the Portrait Value Questionnaire (Schmidt et al. 2007). This PVQ version includes verbal portraits of 21 different people. Each portrait describes a person's goals, aspirations, or wishes that point to the importance of a value. For example: "Thinking up new ideas and being creative is important to her. She likes to do things in her own original way" describes a person for whom self-direction values are important. "It is important to him to be rich. He wants to have a lot of money and expensive things." describes a person who cherishes power values. The respondents' own values are inferred from their self-reported similarity to the portrayed people, i.e. their answers to the following question: "How much like you is this person?", with six labeled responses ranging from "very much like me" to "not like me at all". The scale categories are coded here from 5 to 0 so that 0 stands for "zero importance for me". The importance score of each value is the mean response to the items that measure it. Two or three portraits serve as indicators for ten basic values.

### 2.3 Statistical methods

All data analysis was carried out within the R environment (R Core Team 2016). Our seriousness-of-crimes model was fitted by using the unfolding algorithm of the R-package



**Table 2** Means and standard deviations of crime ratings of samples PF20, MA16, and HD17; persons with NA's omitted

<i>N</i> (persons with complete data)	PF20	MA16	HD17
	1648	2446	2157
Fare evasion	5.34 (1.67)	5.24 (1.69)	4.64 (1.72)
Shop lifting	6.47 (0.97)	6.45 (0.91)	6.24 (1.03)
Drunk driving	6.60 (0.89)	6.59 (0.83)	6.52 (0.87)
Robbery	6.85 (0.53)	6.86 (0.46)	6.81 (0.54)
Cocaine	6.39 (1.24)	6.30 (1.29)	5.93 (1.58)
Smoke pot	5.80 (1.80)	5.51 (1.92)	4.78 (2.15)
Assault	6.87 (0.60)	6.91 (0.43)	6.86 (0.47)
Vandalism	6.58 (0.88)	6.45 (0.97)	6.31 (1.03)
Tax fraud	6.13 (1.32)	6.16 (1.22)	6.08 (1.15)
Benefits fraud	6.56 (0.94)	6.54 (0.86)	6.34 (0.99)
Car break	6.75 (0.74)	6.70 (0.74)	6.51 (0.88)
Taking bribes	6.54 (0.96)	6.58 (0.86)	6.50 (0.90)
Hit & run	6.58 (0.84)	6.48 (0.87)	6.19 (1.03)
Indecent touching	6.77 (0.72)	–	–

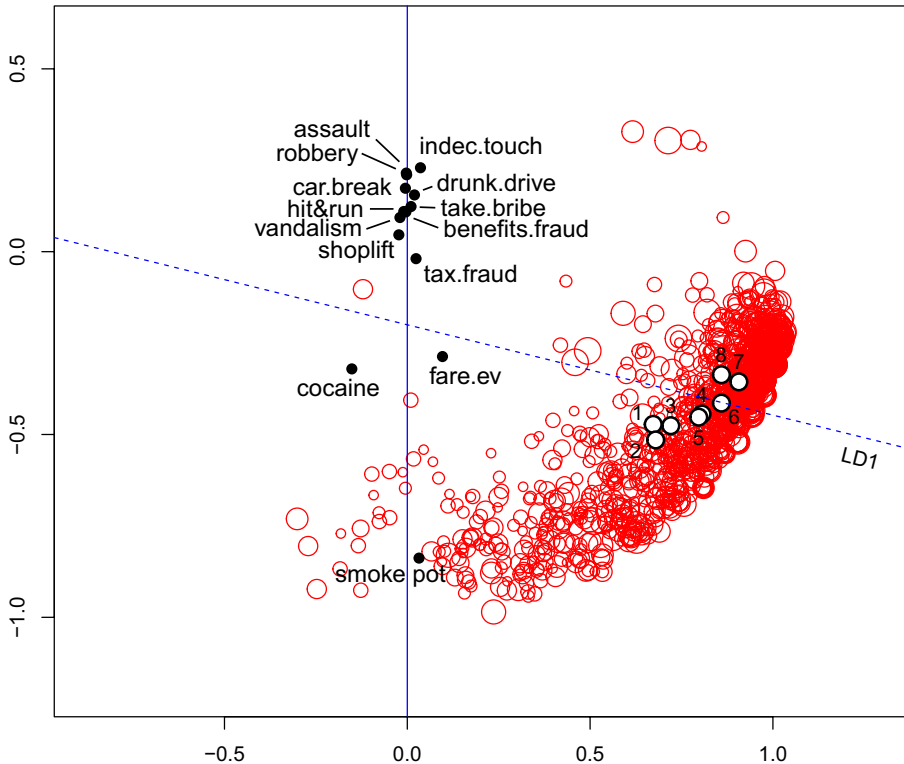
smacof (De Leeuw and Mair 2009; Borg et al. 2018). The distances between a person-point and the points representing the various offenses (the  $d_{ij}$ 's) represent the observed badness ratings of the person (the  $\delta_{ij}$ 's) as precisely as possible in the sense of Stress. We used the strongest scaling algorithm which assumes that the data are ratio-scaled. Thus, no transformations of the data were admitted that formally optimize the loss function but have no substantive theoretical foundation. In the ratio model, the model distances can only be scaled by an overall multiplicative constant which shrinks or enlarges the entire representation space. Ratio-level scaling is the most testable model. It poses the strongest challenge for the hypotheses. It is also the model that is most directly interpretable because it preserves the ratios among the data in the distances. We first test the model fit in 2-dimensional spaces, because this dimensionality is often sufficient to explain psychological proximity data (see Borg and Groenen 2005).

We evaluate the statistical fit of the model to our data by bootstrapping. For each sample, the model fit is tested with 500 random permutations of the seriousness rating data. The distribution of the resulting stress values is then compared with the stress values obtained with the observed data matrices (see Borg et al. 2018; Mair et al. 2016).

The statistical separability of various sub-groups (such as age cohorts) of the person-points in the model solution is tested with discriminant analysis (DA). We used linear DA and the `lda()` function of the MASS R-package for that purpose (Venables and Ripley 2002). Multiple regression was done using the linear model function `lm()` of the R base system (R Core Team 2016).

### 3 Results

Table 2 shows the means and the standard deviations of the 14 (13) items on crimes in the three samples. The mean values show that crimes such robbery, assault, and indecent touching receive the highest seriousness ratings, while fare evasion, smoking pot, and using cocaine are rated as least serious. This is the typical outcome in such surveys, where



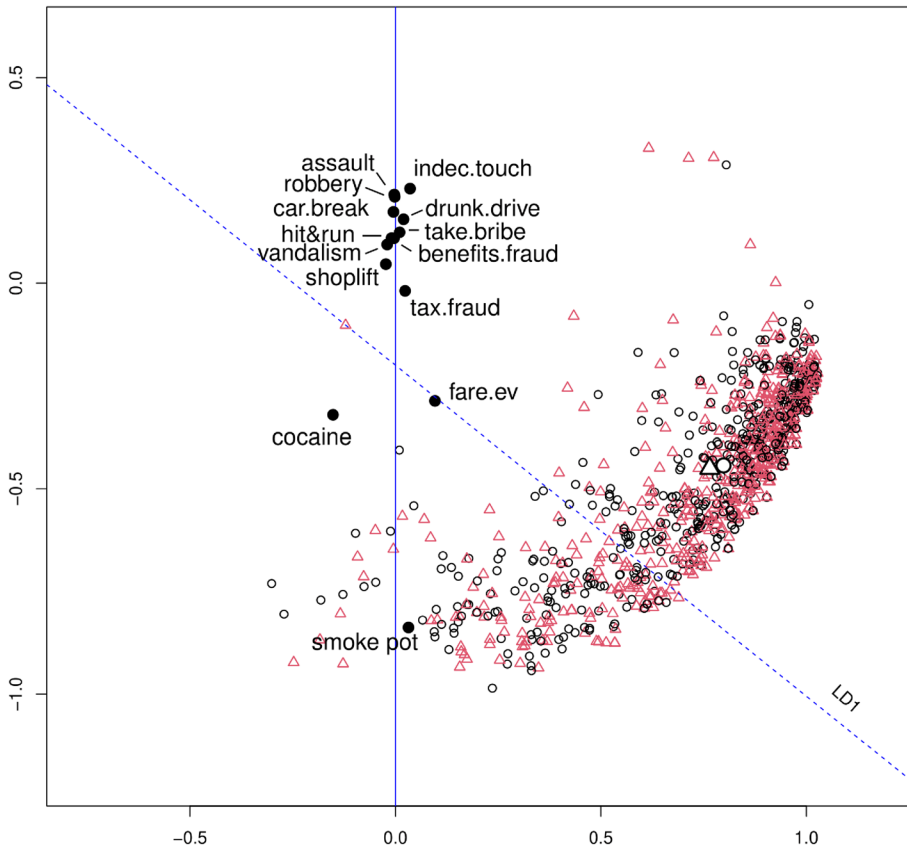
**Fig. 1** Optimal model configuration of seriousness-of-crimes ratings of PF20 sample ( $Stress = .105$ ); rotated to first principal axis of crime points ( $Y$ -axis, vertical line); each open circle represents one of 1,648 persons; size of circles represents persons' age; embedded thick circles show mean person-points of eight age cohorts; dashed blue line (LD1) is first discriminant of age cohorts

the seriousness judgments reflect the wrongfulness and/or the harmful consequences of the offenses. Note that the surveys reported here are representative surveys for three different cities, not for the whole German population.

Turning to the structure of the seriousness ratings, Fig. 1 shows the SOCID model's solution for the PF20 seriousness-of-crimes ratings. It has an excellent and significant fit to the data: The Stress value is .105, with  $p < 0.01$  in the bootstrapping test. Scaling 500 random permutations of the PF20 data with 2d ratio unfolding yields a 1% percentile Stress value of 0.160 ( $mean = .161$ ;  $min = .156$ ). Hence, the SOCID model succeeds in explaining the PF20 sample data by a 2-dimensional Euclidean map that represents the 14 crimes by 14 points and each of the respondents by one point per person<sup>3</sup>.

The plot can be arbitrarily rotated because rotations do not change the distances among the points. Figure 1 exhibits a rotation that unclutters the labels of the crime points as

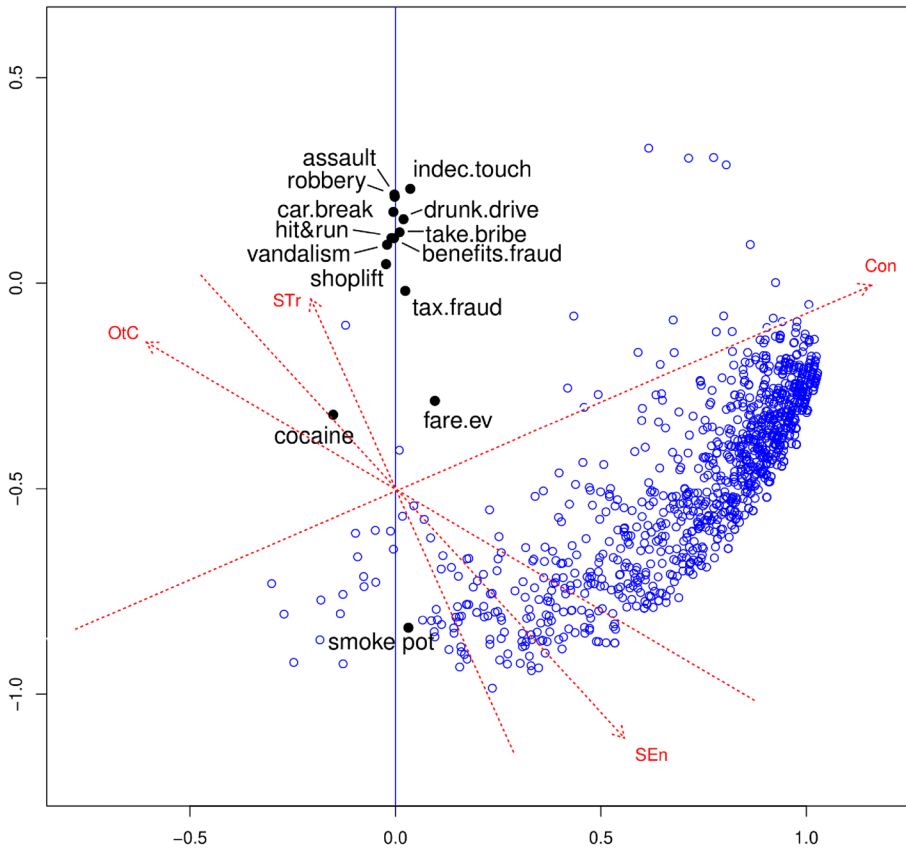
<sup>3</sup> If one runs weaker versions of the model, assuming only interval or ordinal scale level of the data, one obtains very similar results. The Stress values found by *smacof* unfolding are .124 and .123, respectively. The configurations tend to cluster the crime points somewhat more than the ratio-scale model solution in Fig. 1, while the person points exhibit the same horseshoe-like pattern as in Fig. 1 in both of the weaker model solutions.



**Fig. 2** Optimal model configuration as in Fig. 1; triangles represent men, circles women; embedded thick triangles/circles show centroids of males and females, respectively; dashed blue line (LD1) is first discriminant for gender

much as possible. The Y-axis of the plot is the first principal axis of the crime points, and the X-axis is the second principal axis. The vertical line running through the crime points shows that pot-smoking is at the bottom end and indecent touching at the top end of the first principal axis. This confirms the hypothesis that the model will uncover a common crime point configuration that is essentially linear and that can be interpreted in the usual way (wrongfulness and/or harmful consequences).

Figure 1 also shows how persons who belong to one of eight age cohorts are distributed in model space: The older they are, the larger the circles that represent the persons. The eight thick black circles in the cloud of the person-points represent the centroids of the persons in the age cohorts 1, 2, ..., 8. One notes that larger circles tend to be positioned farther away from the crime scale. This impression is supported using linear discriminant



**Fig. 3** Optimal model configuration as in Fig. 1; circles represent persons; four dashed oriented lines are scales estimating the persons' observed scores on the higher-order values Con, SEn, STr, and OtC

analysis (LDA). The dashed blue line in Fig. 1 shows the first discriminant (LD1): Projecting the person-points onto this line separates the eight age cohorts optimally. The proportion of the trace generated by the LD1 is .816. The correlation of the point projections with the persons' age category is  $r = .286$  ( $t = 10.7$ ,  $N = 1284$ ,  $p < .01$ , with a 95% confidence interval of [0.24, 0.34]). This geometrically expresses that the older a person, the more he or she tends to rate *all* crimes as relatively serious. Younger persons, in contrast, rate the crimes more differently and petty offenses (such as pot-smoking) as not so serious<sup>4</sup>.

Figure 2 shows the same configuration as Fig. 1, but here the symbols describing the person-points distinguish between male (triangles) and female (circles) persons. The plot suggests that the genders are thoroughly mixed in the distribution of person-points. That is, the points cannot even roughly be partitioned into sub-spaces containing only male or female respondents. LDA nevertheless identifies an optimal discriminant,

<sup>4</sup> Multiple regression rather than LDA, predicting the persons' age by the  $X$ - and  $Y$ -coordinates of the person-points, leads to almost the same projection line as LDA, with a correlation of the observed age scores and the regression-based age predictions of  $r = .287$ .

**Table 3** Stress values of SOCID representations of persons of samples PF20, MA16, and HD17, resp.; correlations of mean seriousness-of-crimes (SoC) ratings with external variables (four personal values, age, gender; “+” for multiple regression)

	PF20	MA16	HD17
<i>N</i> (persons with complete data)	1648	2446	2157
Stress of ratio-unfolding solution	0.105**	0.108**	0.117**
<i>r</i> (mean(SoC ratings), self-enhancement )	-.10**	-.18**	-.18**
<i>r</i> (mean(SoC ratings), self-transcendence)	.06**	.02	.00
<i>r</i> (mean(SoC ratings), openness to change)	-.15**	-.22**	-.21**
<i>r</i> (mean(SoC ratings), conservation)	.28**	.37**	.38**
<i>r</i> (mean(SoC ratings), age)	.29**	.37**	.39**
<i>r</i> (mean(SoC ratings), age + conservation)	.38**	.49**	.51**
<i>r</i> (mean(SoC ratings), gender)	.08**	.08**	.10**

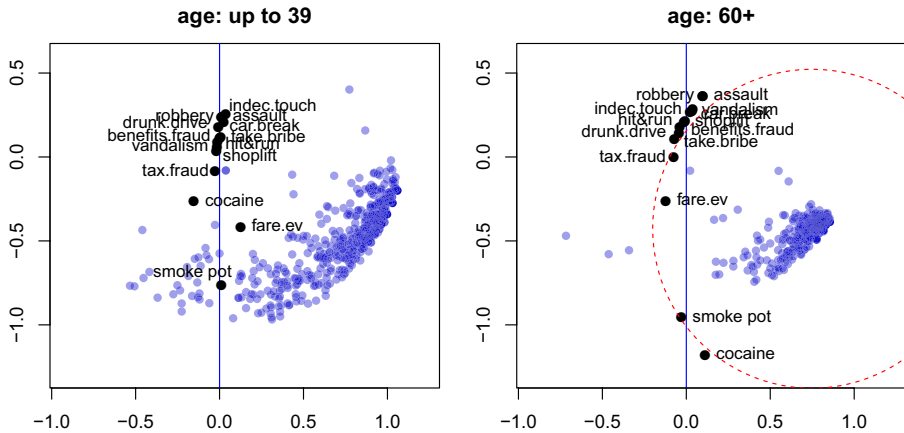
**Table 4** Inter-correlations of the coordinates of the crime points on the first principal axes of the SOCID solutions of three survey samples (see vertical blue line in Fig. 1)

	PF20	MA16	HD17
PF20	1	.84	.98
MA16	.84	1	.88
HD17	.98	.98	1

shown in Fig. 2 by the blue dashed line. The correlations of the projections of the person-points onto this line with their observed gender categorization is, however, only  $r = .087$  ( $t = 3.13$ ,  $df = 1281$ ,  $p$ -value = 0.002, with a 95% confidence interval of [0.03, 0.14]). Hence, the effect of gender on the distribution of the person-points is negligible (Table 3).

Finally, Fig. 3 shows how the distribution of the person-points is related to the persons’ striving for four different higher-order personal values. The plots also show four dashed lines. Each line represents a scale computed by estimating the observed higher-order personal value ratings of persons  $p = 1, \dots, N$  ( $obs_p$ ) by multiple regression on the  $X$  and  $Y$  coordinates of the person points:  $obs_p \approx a + w_1 * X_p + w_2 * Y_p$ , with weights  $a, w_p$ , over all persons  $p$ . Each higher-order personal value is significantly related to the SOCID coordinates of the persons, but only Conservation (aggregating a person’s striving for tradition, conformity, and security values) is substantial in size. The correlations of the observed higher-order personal value scores and the regression-based estimates are .117\*\* (for self-enhancement, SEn), .159\*\* (for openness to change, OtC), .111\*\* (for self-transcendence, STr), and .287\*\* (for conservation, Con). For Conservation, the data show that persons who more strongly strive for this higher-order personal value tend to be farther from the crime scale than persons who are less oriented in this value direction. Hence, conservation has a similar effect (in its West-East orientation) on the persons’ positions in scaling space as age.

Value research (Schwartz 1992) has shown that persons who find self-enhancement values (SEn: power, achievement) important rate self-transcendence values (STr: benevolence, universalism, self-direction) as less important, and vice versa. A similar opposition holds for conversation (Con: tradition, conformity) and openness to change (Con: hedonism,



**Fig. 4** Optimal model configurations of seriousness-of-crime ratings for young (up to 39 years) and old (60 years and older) persons of PF20 sample ( $Stress = .127$  and  $.086$ , resp.); center of circle in plot on the right-hand side is centroid of person-points

stimulation). This is clearly reflected by the regression lines in Fig. 3 for the pair *STr* and *SEn*, and also roughly true for *Con* and *OtC*.

The trends in the person-points' positions related to age, gender, and conservation reflect what can also be found directly in the data. Table 4 shows, for example, that the higher-order value Conservation correlates significantly and substantially with the persons' mean seriousness ratings. Moreover, it also correlates significantly and substantially but negatively with the higher-order value that is theoretically opposed to Conservation, i.e. with Openness to Change. The correlations of age with the average seriousness scores are similar in size, but the effects are not equivalent: When combining age and Conservation in multiple regression, the correlations are even larger. Gender is essentially not correlated with the seriousness ratings. This is also evident from Fig. 2.

The *SOCID* representations of the samples *MA16* and *HD17* are not shown here. They are very similar to the configuration in Fig. 1. In particular, the configuration of the crime points is again almost linear and it exhibits the wrongfulness and consequences gradients. This can be shown numerically by correlating the first-principal-axis coordinates of the crime points of the three solutions (Table 4).

Figure 4 shows the scaling solutions for two subgroups of the PF20 data, persons under 40 years of age ( $Stress = .127$ ) and persons aged 60 or older ( $Stress = .086$ ). The younger persons' solution is similar to the solution of the total sample. The person-points scatter extensively, and the crime points replicate the seriousness scale of Fig. 1. In the solution of the older persons, in contrast, the person-points are clustered much more densely. Also, cocaine has moved down on the seriousness scale to a bottom position. The circle about the centroid of the person-points in the right panel of Fig. 4 makes clear that most distances from person-points to crime points are quite similar. Cocaine, therefore, could also be moved upwards along the circle without much effect on the Stress of the solution. Hence, the solution is not very robust in this case where most persons exhibit relatively similar ratings: The mean seriousness rating

of the young persons is 6.17 ( $sd=0.81$ ) and 6.65 ( $sd=0.52$ ) for the older persons (on the 7-point rating scale).

## 4 Discussion

It could be shown that the respondents of three large and representative surveys on crime prevention in different German cities generated seriousness-of-crime ratings that can be explained both formally and substantively by a geometric distance model (SOCID). The model says that all persons share a *common* seriousness scale. The scale we found is essentially uni-dimensional and resembles the seriousness scales typically reported in the literature, ranging from petty offenses to serious crimes. In contrast to previous research, this scale is not simply an automatic product of scale construction such as simply computing mean ratings (as in Table 2). Rather, the scale is found embedded in a 2-dimensional plane, where every individual is represented by a single point that is positioned such that its distances to the crime points on the seriousness scale correspond closely to the observed ratings of each individual. The essentially linear ordering of the crime points in Figs. 1, 2, and 3 results from the constraints contained *in the data* and not from formal constraints of the model. If, for example, crimes  $X$  and  $Y$  are positioned close to each other in crime space, then the distances from *any* person-point in crime space will be relatively similar, no matter how high the dimensionality ( $m$ ) of the space. If there are persons in the sample who perceive  $X$  and  $Y$  as relatively different, then these persons cannot be properly represented in a crime space where the crime points of  $X$  and  $Y$  are close neighbors. Thus, *both* crime points and person-points are iteratively moved in  $m$ -dimensional space by the scaling algorithm until a joint configuration with minimal Stress has been found. Any configuration of crime points and person-points is possible, in principle, as a SOCID solution: The algorithm minimizes only the Stress function, given the dimensionality  $m$ . Thus, a solution with an acceptably small Stress that also makes substantive sense is a remarkable finding.

The seriousness ratings are interpreted as judgments of distance: The higher the person's seriousness rating, the more the person distances him- or herself from the crime. This distancing is systematically related to the persons' age and their orientation towards such personal values as tradition, conformity, stimulation, etc. The older and the more conservation-oriented the person, the more distant he/she is from the scale in model space. This corresponds to the means of the rating data. The persons' striving for a particular personal value, conservation (i.e., tradition, conformity, and security) has a similar effect on the distribution of the person-points in the psychological map. Gender does not allow discriminating the person-points on the map.

The reported model solutions were all found assuming that the observed seriousness ratings are on the level of a ratio scale. Moreover, we restricted the dimensions of the solution space to only two. If the dimensionality of the model space is enlarged, the Stress drops from 0.105 for  $m = 2$ , to 0.090 for  $m = 3$ , to 0.082 for  $m = 4$ . Thus, enlarging the solution space does not explain substantially more variance. A 2-dimensional crime space is sufficiently precise. The excellent fit explaining the seriousness ratings of all 1,648 individuals of the PF20 sample, is remarkable.

A particularly interesting property of our scaling solutions is the essential uni-dimensionality of the crime points in the scaling solutions. This finding should, however, be taken with some care. First, the distribution of the 14 crime points is quite skewed. Most crimes are clustered at one end of the scale, and only a few petty crimes stretch

out the scale at the bottom end. Hence, what remains to be shown in future research is whether such uni-dimensional scales can also be found if one uses *more* crimes, and crimes with a larger *range* of seriousness. Second, the crimes in our surveys do not exhibit the factors of a systematic construction design as it is often the case in studies that elaborate the crime scenarios in much detail. Hence, in an anonymous survey where more or less spontaneous answers are sufficient, the respondents may utilize a *simplified* mental map that is indeed dominated by a stereotypical ordering of crimes in terms of wrongfulness and/or harm. Third, in a data collection format where the respondents are asked to also rate the crimes on other scales, and that even before the seriousness ratings (as in Adriaenssen et al. 2017), one may need a crime space with more than two dimensions to represent the data with low Stress. It also seems likely that the crime points would not arrange themselves in a higher-dimensional space in form of a simple linear manifold. Indeed, Adriaenssen et al. (2017) showed that wrongfulness is the best predictor of seriousness, but harm is a good predictor too. Hence, if the individuals are prompted to make more analytical judgments, it seems likely that they use additional criteria for their seriousness assessments, and this *may* lead to a more complex crime point configuration.

From a practical and applied perspective, the findings reported in this paper can be seen as reassuring that scales derived by simply averaging seriousness ratings (as in Table 1), without any testing of the scalability of the items, are reliable and valid – at least for items and samples similar to those utilized in this study. This is certainly a non-trivial result, because the structure of seriousness-of-crimes ratings could also be multi-dimensional, for example, with different sub-scales for different types of crimes (e.g., property vs. violence crimes). In any case, given a particular sample and seriousness-of-crimes ratings, it is always possible to test the scalability of the data using the SOCID approach. This information cannot be derived using the usual across-person analyses: The results they generate may be artifacts, mixing different types of persons and generating, thereby, potentially meaningless scales of the “average person” that may not describe most individuals.

Another issue of applied interest is that one can test how much each item or each person contributes to the overall stress of the model, thereby identifying outliers or sub-sets of items/persons that should be studied more closely. One simply computes  $\sigma$  in formula (1) for each item or person separately,

$$\sigma_j = \sum_{i=1}^N (\delta_{ij} - d_{ij})^2, \text{ for item } j, \quad (4)$$

$$\sigma_i = \sum_{j=1}^n (\delta_{ij} - d_{ij})^2, \text{ for person } i, \quad (5)$$

and then normalizes the result by dividing it by the sum of all  $\sigma_j$ 's or  $\sigma_i$ 's. This yields a *Stress-per-point* (SPP) measure that shows how much each offense  $j$  or each person  $i$  contributes to the global Stress.

For the 14 offenses in Fig. 1, each crime is expected to contribute  $1/14 \cdot 100\% \approx 7\%$  to the global Stress. Applying formula (4) to the PF20 data and its SOCID distances shows that fare evasion contributes 19% and tax fraud 14%, while robbery contributes only 3% and smoking pot 4%. The relatively poor fit of fare evasion and its somewhat unexpected (see its mean value in Table 2) position in the middle of the seriousness scale of the SOCID solution means that the fare-evasion ratings are relatively inconsistent with the rest



of the ratings. One possible explanation for this observation is that many respondents were ambiguous about this offense. Another explanation is that they considered fare evasion as not truly comparable to the other offenses. Further research may ask the respondents to explain how they arrived at their ratings, and how certain they felt about their scoring.

The SPP values of the various persons in the sample exhibit an extremely skewed distribution, with most persons fitting well into the solution in Fig. 1. Persons who do not fit into the solution are easily detected by their SPP values. Their ratings can then be studied more closely. For example, the person with the highest SPP value rates all offenses in Table 1 as very serious—except for the last four offenses (car break, etc.) that are rated as not serious at all. In this case, no rule that would generate such judgments is obvious. In contrast, there is a group of persons with high SPP values that produced ratings consistent with the solution in Fig. 1, except that they rate all property-related crimes as not so serious.

Some problems of the scaling model as the mathematical framework for robust scaling are also clear: If there is little variance in the data, the unfolding algorithm tends to generate solutions with excellent fit (i.e., low Stress) but *degenerate* configurations. We noticed in Fig. 4 that the solution for the old age cohort (which rates all crimes as quite serious) approximates a degenerate solution where all person-points sit in the center of a circle of crime points. In a perfectly degenerated solution, the person-points can be moved arbitrarily on this circle without affecting the Stress. Alternatively, all crime points could be positioned at a circle's center and all person-points onto the circle.

The above problem is a consequence of the data collection method and also of the particular collection of crimes used as items. The items in Table 1 generate rather skewed distributions of seriousness ratings, where only a few items receive low seriousness ratings but most items get high scores. Indeed, 16% of the respondents rated all crimes the same, with 98.6% of them scoring all crimes as “very serious”. This ceiling effect must always be taken into account when collecting seriousness-of-crime ratings. One method that may at least reduce the effect is adding a truly serious crime (such as “murder”) to the set of crimes. This would establish a frame of reference that pushes the ratings of the other crimes to a lower level. Another approach is using different scaling methods such as rankings instead of ratings, magnitude scaling, or pair comparisons, but such methods are generally too time-consuming and too demanding for typical surveys.

Finally, concerning applied consequences of our findings, one can argue that they have practical relevance for the justice system, especially with regard to sentencing. Sentencing guidelines emanate from the following section of the German Criminal Code: §46 StGB. The offender's guilt provides the basis on which the penalty is fixed. This includes the guilt of the criminal act and the guilt of the conduct of life. The English guidelines reflect a proportional sentencing model, oriented on the seriousness of crime (Roberts and Padfield 2020; Jehle 2020). In criminal law, the principle of proportional justice is used to describe the idea that the punishment of a certain crime should be in proportion to the severity of the crime itself. The fundamental principle behind proportionality is that the punishment should fit the crime (Hörnle 1999). Since law does not define the severity of offenses, it should be helpful for judges to have information about people's ratings of the seriousness of crimes.

**Author Contributions** Conceptualization, methodology, formal analysis and investigation, writing: IB, Data: DieterH, Writing—review and editing: WB.

**Funding** Open Access funding enabled and organized by Projekt DEAL. None.

## Declarations

**Conflict of interest** The author declares that they have no conflict of interest.

**Data availability** The data belong to the cities who financed the crime surveys; for possible scientific use, contact Dieter Hermann.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abrams, J.K., Della-Fave, R.L.: Authoritarianism, religiosity, and the legalization of victimless crimes. *Soc. Soc. Res.* **61**, 68–82 (1976)
- Adriaenssens, A., Paoli, L., Karstedt, S., Visschers, J., Greenfield, V.A., Pleysier, S.: Public perceptions of the seriousness of crime: weighing the harm and the wrong. *Eur. J. Criminol.* **17**(2), 127–150 (2017). <https://doi.org/10.1177/1477370818772768>
- Allerbeck, K.R., Mayer, K.U., Müller, W.: Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 1990 [General population survey in the social sciences] (Tech. Rep.). Köln, Germany: GESIS Datenarchiv. (2017). <https://doi.org/10.4232/1.1800>
- Bentrop, C., Schmid, J.B., Tzvekkova, M., Vreden, C.J.: Codebuch der kriminologischen Schülerbefragung in Münster (2000–2003) [Codebook of the criminological student survey in Münster (2000–2003)] (Tech. Rep.). Bielefeld, Germany: Uni Bielefeld. (2017). Retrieved from <https://uni-bielefeld.de/fakultaeten/soziologie/forschung/projekte/krimstadt/pdf/CodebuchMS.pdf>
- Beullens, K., Matsuo, H., Loosveldt, G., Vandenplas, C. (Eds.): Quality report for the European Social Survey, Round 6. London, UK: European Social Survey ERIC. (2014) Retrieved from <http://www.europeansocialsurvey.org/docs/round6/methods/ESS6qualityreport.pdf>
- Bilsky, W., Borg, I., Hermann, D.: Utilizing personal values to explain persons' attitudes towards legal norms. *Eur. J. Criminol.* (2020). <https://doi.org/10.1177/1477370820913237>
- Borg, I.: Age and the subjective importance of personal values. *Pers. Individ. Differ.* (2021). <https://doi.org/10.1016/j.paid.2020.110605>
- Borg, I., Bardi, A., Schwartz, S.H.: Does the value circle exist within persons or only across persons? *J. Pers.* **85**, 151–162 (2017). <https://doi.org/10.1111/jopy.12228>
- Borg, I., Groenen, P.J.F.: Modern multidimensional scaling (2nd ed.). New York: Springer (2005). <https://doi.org/10.1007/0-387-28981-X>
- Borg, I., Groenen, P.J.F., Mair, P.: Applied Multidimensional Scaling and Unfolding (2nd ed.). New York: Springer (2018). <https://doi.org/10.1007/978-3-642-31848-1>
- Borg, I., Hermann, D.: Personal values of lawbreakers. *Pers. Individ. Differ.* (2020). <https://doi.org/10.1016/j.paid.2020.110104>
- Borg, I., Hermann, D.: Inside and outside perspectives on the relation of people's personal values and their acceptance of legal norms. *Macro Management & Public Policies.* **03**(04), 1–13 (2021). <https://doi.org/10.30564/mmpp.v3i4.4178>
- Borg, I., Staufenbiel, T.: Theorien und Methoden der Skalierung [Theories and methods of scaling], 4th edn. Huber, Bern, Switzerland (2007)
- Busing, F.M., Heiser, W.J., Cleaver, G.: Restricted unfolding: preference analysis with optimal transformations of preferences and attributes. *Food Qual. Prefer.* **21**(1), 82–92 (2010). <https://doi.org/10.1016/j.foodqual.2009.08.006>

- Carroll, J.D., Chang, J.-J.: Analysis of individual differences in multidimensional scaling via an n-way generalization of eckart-young decomposition. *Psychometrika* **35**, 283–319 (1970). <https://doi.org/10.1007/BF02310791>
- Coombs, C.H.: *A Theory of Data*. Wiley, New York (1964)
- De Leeuw, J., Mair, P.: Multidimensional scaling using majorization: SMACOF in R. *Journal of Statistical Software*, **31**(3), 1–30 (2009). <https://doi.org/10.18637/jss.v031.i03>
- Dietrich, F.: Recht als Sonderfall einer Normordnung [Law as a special case of a normative system]. In E. Hilgendorf & J. Joerden (Eds.), *Handbuch Rechtsphilosophie* (pp. 2–7). Stuttgart: Metzler (2017)
- Dunn-Rankin, P., Knezek, G.A., Wallace, S., Zhang, S.: *Scaling Methods*. Lawrence Erlbaum, Mahwah, New Jersey (2004)
- Gigerenzer, G., Gaissmaier, W.: Heuristic decision making. *Ann. Rev. Psychol.* **62**(1), 451–482 (2011). <https://doi.org/10.1146/annurev-psych-120709-145346>
- Hörnle, T.: *Tatproportionale Strafzumessung [Offense-proportional sentencing]*. Duncker & Humboldt, Berlin (1999)
- Hermann, D.: *Werte und Kriminalität: Konzeption einer allgemeinen Kriminalitätstheorie [Conception of a general theory of criminality]*. Westdeutscher Verlag, Wiesbaden, Germany (2003)
- Hermann, D.: *Heidelberger Sicherheitsaudit 2017 [Security audit 2017 in Heidelberg]* (Tech. Rep.). Heidelberg, Germany: Universität Heidelberg, Institut für Kriminologie (2017a)
- Hermann, D.: *Mannheimer Sicherheitsaudit 2017 [Security audit 2017 in Mannheim]* (Tech. Rep.). Heidelberg, Germany: Universität Heidelberg, Institut für Kriminologie (2017b)
- Hermann, D., Wachter, E.: *Pforzheimer Sicherheitsaudit 2020 [Security audit 2020 in Pforzheim]* (Tech. Rep.). Heidelberg, Germany: Universität Heidelberg, Institut für Kriminologie (2017)
- Hirtenlehner, H., Reinecke, J.: Moralische Grenzen der instrumentellen Verhaltenssteuerung - Hängt die Bedeutung der Sanktionsrisikoeinschätzung vom Ausmass innerer Normbindung ab? [Moral bounds of instrumental behavioral regulation - does the importance of expected sanctions depend on norm acceptance?]. *Monatsschrift für Kriminologie und Strafrechtsreform* **101**, 380–402 (2018). <https://doi.org/10.1515/mks-2018-1013-409>
- Jehle, J.-M.: *Sentencing in England and Wales. a commentary from a German perspective*. In K. Ambos (Ed.), *Sentencing: Anglo-American and German insights* (p. 91–99). Göttingen, Germany: Göttingen University Press (2020). <https://doi.org/10.17875/gup2020-1330>
- Jowell, R., Roberts, C., Fitzgerald, R., Eva, G. (Eds.): *Measuring attitudes crossnationally: Lessons from the European Social Survey*. London, UK: Sage (2007). Retrieved from <https://psycnet.apa.org/record/2015-15039-000>
- Mair, P., Borg, I., Rusch, T.: Goodness-of-fit assessment in multidimensional scaling and unfolding. *Multivar. Behav. Res.* **51**, 772–789 (2016). <https://doi.org/10.1080/00273171.2016.1235966>
- McCleary, R., O'Neil, M.J., Epperlein, T., Jones, C., Gray, R.H.: The context of normative consensus: an expansion of the rossi/berk consensus model with an application to crime seriousness. *Soc. Sci. Res.* **20**, 1–28 (1991). [https://doi.org/10.1016/0049-089X\(91\)90001-J](https://doi.org/10.1016/0049-089X(91)90001-J)
- McCleary, R., O'Neil, M., Epperlein, T., Jones, C., Gray, R.H.: Effects of legal education and work experience on perceptions of crime seriousness. *Soc. Problems* **28**, 276–289 (1981). <https://doi.org/10.2307/800303>
- R Core Team.: *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria (2016). Retrieved from <https://www.R-project.org/>
- Roberts, J.V., Padfield, N.: *Sentencing in England and Wales*. In K. Ambos (Ed.), *Sentencing: Anglo-American and German insights* (p. 71–90). Göttingen, Germany: Göttingen University Press (2020). <https://doi.org/10.17875/gup2020-1330>
- Rossi, P.H., Waite, E., Bose, C.E., Berk, R.E.: The seriousness of crime: normative structure and individual differences. *Am. Sociol. Rev.* **39**, 224–242 (1991). <https://doi.org/10.2307/2094234>
- Schmidt, P., Bamberg, S., Davidov, E., Herrmann, J., Schwartz, S.H.: Die Messung von Werten mit dem Portraits Value Questionnaire. *Zeitschrift für Sozialpsychologie* **38**, 261–275 (2007). <https://doi.org/10.1024/0044-3514.38.4.261>
- Schwartz, S.H.: Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–65) (1992). [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6)
- Schwarz, N.: The psychology of survey response. In: W. Donsbach & M. W. Traugott (Eds.), *The SAGE handbook of public opinion research* (pp. 374–387) (2008). <https://doi.org/10.4135/9781848607910.n35>

- Seddig, D.: Peer group association, the acceptance of norms and violent behaviour: a longitudinal analysis of reciprocal effects. *Eur. J. Criminol.* **11**, 319–339 (2014). <https://doi.org/10.1177/1477370813496704>
- Seddig, D.: Crime-inhibiting, interactional and co-fevelopmental patterns of school bonds and the acceptance of legal norms. *Crime Delinq.* **62**(8), 1046–1071 (2016). <https://doi.org/10.1177/0011128715578503>
- Sellin, T., Wolfgang, M.: *The Measurement of Delinquency*. Petterson Smith, Montclair (1964)
- Sellin, T., Wolfgang, M.: *The National Survey of Crime Severity*. U.S. Government Printing Office, Washington, DC (1985)
- Stylianou, S.: Measuring crime seriousness perceptions: what have we learned and what else do we want to know. *J. Crim. Justice* **31**(1), 37–56 (2003). [https://doi.org/10.1016/S0047-2352\(02\)00198-8](https://doi.org/10.1016/S0047-2352(02)00198-8)
- Tourangeau, R., Rips, L.J., Rasinski, K.: *The Psychology of Survey Response*. Cambridge University Press, Cambridge (2003). <https://doi.org/10.1017/CBO9780511819322>
- Venables, W.N., Ripley, B.D.: *Modern applied statistics with s* (Fourth ed.). New York: Springer (2002). Retrieved from <https://www.stats.ox.ac.uk/pub/MASS4/> (ISBN 0-387-95457-0)
- Walker, M.A.: Measuring the seriousness of crimes. *Criminology* **18**(4), 348–364 (1978). <https://doi.org/10.1093/oxfordjournals.bjc.a046933>
- Warr, M.: What is the perceived seriousness of crimes. *Criminology* **27**(4), 795–814 (1989). <https://doi.org/10.1111/j.1745-9125.1989.tb01055.x>
- Wasmer, M., Koch, A., Wiedenbeck, M.: *Methodenbericht zur 'Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften' ALLBUS 1990* [Methods report of the general population survey in the social sciences] (Tech. Rep.). Mannheim, Germany: ZUMA (1991). Retrieved from <https://www.gesis.org/allbus/inhalte-suche/methodenberichte>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.