İSTANBUL
UNIVERSITY
PRESS

Araştırma Makalesi / Research Article

What Makes a Crime Serious? Testing Warr's Model of Offence Seriousness

Bir Suçu Ağır Yapan Nedir? Warr'un Suç Ağırlığı Modelinin bir Testi

Galma AKDENİZ[1]

[1]Dr. Öğr. Üyesi, İstanbul Bilgi Üniversitesi, Hukuk Fakültesi, İstanbul, Türkiye

ORCID: G.A. 0000-0002-7255-8260

ABSTRACT

Warr (1989) proposed that the perceptions of seriousness of different criminal offences are a function of perceptions of harmfulness caused by a crime (consequences of the crime), and perceived wrongfulness of a crime (normative evaluations regarding the crime). The study reported in this paper tested this model and examined the perceptions of seriousness of different offences in a sample of university students in Turkey. It was found that the degree of consensus regarding offence seriousness was much higher for offences judged as more serious. It was further found, when using wrongfulness and harmfulness assessments, that offences clustered into three larger groups: offences that present threat/risk of physical harm/death, property offences, and "minor" offences. Further, Warr's model was tested on both the offence and the individual level of analysis. The findings suggest that the model indeed holds, however relative strength of harmfulness and wrongfulness, as predictors of crime seriousness, are different in Turkey, compared to findings from the USA and Westernest European countries. On an individual level, it was found that harmfulness was a stronger predictor than wrongfulness in a much larger number of offences, and that on the level of the offence, harmfulness was as strong a predictor of seriousness as wrongfulness.

Keywords: Crime seriousness, criminal victimization, norm violation, offences, violence perceptions

Citation: Akdeniz G, 'What Makes a Crime Serious? Testing Warr's Model of Offence Seriousness' (2020) 8(1) Ceza Hukuku ve Kriminoloji Dergisi-Journal of Penal Law and Criminology, 1.

1. Introduction

Classification of crime by seriousness is fundamental to how criminal policy, and criminal law as its tool, is constructed today. The just desserts and the retributive approaches to criminal sanctions require longer sentences, due to proportionality principle, for more serious crimes. Such crimes are considered to present a more serious threat to public safety and to cause greater harm to the society and therefore are seen as needing more urgent and extensive intervention when it comes to prevention and policy (Miethe, 1982; Rose & Prell, 1955; Rossi, Waite, Bose, & Berk, 1974; Cullen, Link, Travis, & Wozniak, 1985).

Proportionality of punishment to the crime is considered one of the fundamental principles of penal law in Turkey. This principle has been embodied into the law, with Art. 3(1) of the Turkish Criminal Code which states that "The offender is to be given a sentence […] that is proportionate to the seriousness of the offences committed". The Turkish Criminal Code specifies a range of possible sentences for each offence. This determines a starting point for calculating the sentence in any particular case, with different factors, identified in the Criminal Code in Articles 61-63, being used in calculating the sentence on case-by-case basis. Consequently, while the proportionality is far from being the sole determinant of the sentencing decisions in Turkish criminal justice system, it can be concluded that the principle of proportionality utilizing offence seriousness is a basic pillar of the Turkish penal law's punitive structure.[1]

However, how exactly sentence ranges specified in the Turkish Criminal Code were determined, or how one can be sure that they are actually proportionate to the seriousness of the offence, is not obvious. Even further, what "seriousness" is and what determines it is unclear. It is often presumed that the punishments prescribed in the penal code are in fact proportionate to the seriousness of crimes, and that in individual instances the judges are able to determine what exact sentence is proportional to the seriousness of crime in that particular instance. However, when the Turkish Criminal Code was written and adopted, there was no real discussion regarding this issue, in either the political or academic realm. In other words, what crimes are more or less serious, and therefore which crimes deserve more or less serious punishment is not based on any particular systematic analysis, other than perhaps legal precedent, tradition, and what appears to be common sense. But, is there really a "common" sense regarding crime

---

1   See Üzülmez (2006) and Taneri (2006) for discussion of this topic.

seriousness, reflected as general agreement regarding seriousness levels of different crimes? This question of how to quantify the seriousness of crime, along with the question of which characteristics of an offence make it more or less serious, is a broad subject addressed by the research discussed in this paper.

The proportionality of punishment is not the only reason why understanding and ability to measure (and therefore quantify) seriousness is important. Another area where the ability to quantify the seriousness of offences would be of great use is the identification of crime trends. Crime trends are often identified by looking at frequencies of different offenses (usually a subset of so called "index crimes", such as burglary, robbery, theft, assault, sexual assault, etc.), but the fact that not all crimes are "equal" in terms of their individual contribution to the size of the safety problem presents an issue: Crimes that are perceived as more serious contribute more to our perception of the extent of the crime problem. Hence, a large number of less serious crimes may be perceived as less of a problem than a smaller number of crimes that are perceived as more serious. Developing a way to quantify the contribution of each type of offense to the crime problem in a way that is more sophisticated than the mere frequency count, would allow governments to follow and track the size of crime problem in greater detail (Burton, Finn, Livingston, Scully, Bales, & Padgett, 2004; Kwan, Ip, & Kwan, 2000; Sebba, 1980). Hence, for example, a trend that involves a drop in homicide numbers accompanied by an identical increase in the number of simple thefts would not reflect as a change if measured by simple frequencies. However, if one was to quantify the seriousness of offences, then the trend could reflect this quantification, thus detecting what in fact could be an increase in the size of a crime problem. While some research results suggest that such an approach to measuring crime trends may not be particularly useful, Stylianou has argued that studying how increases in perceived seriousness of particular offences contribute to the public's overall assessments of crime problem can be of great use (2003, p. 49).

One of the first attempts to define and measure crime seriousness was by Sellin and Wolfgang (1964). Since, criminologists have worked to further improve their method, based on the framework that defines the issue as a question of individual judgment (Carlson & Williams, 1993; Fishman, Kraus, & Cohen, 1986; O'Connell & Wheelan, 1996). Over time, in criminological research the idea that *offence seriousness* is not an intrinsic or objective quality of the criminal offence itself, but rather a reflection of a *judgment* regarding crime by regular citizens has become an accepted norm. Which offence is more and which is less serious, therefore, can only be quantified by asking

common citizens to rate, rank, or compare different offenses by how "serious" they judge them to be. From this point, the research has developed asking two distinct questions: 1) what characteristics of an offence lead it to be perceived as more/less serious; 2) what is behind individual variability in offence seriousness assessment.

 Two main issues that research had focused on are, on one hand, the degree of consensus regarding crime seriousness judgments both within societies, as well as cross-culturally (Lesieur & Lehman, 1975; Cullen, Link, & Polanzi, 1982; Parton, Hansel, & Stratton, 1991; Levi & Jones, 1985; Fishman, Kraus, & Cohen, 1986; Kwan & Kwan, 2000; Herzog & Einat, 2016). While research shows that there is a substantive degree of consensus regarding the seriousness of the most and the least serious crimes, there is somewhat less agreement when it comes to crimes which fall in the midrange (Cullen, Link, Travis, & Wozniak, 1985). Large research projects, such as that conducted by the U.S. Bureau of Justice Statistics in 1985, which covered over 200 different offenses and included 50.000 participants have attempted to address this issue (Wolfgang, Figlio, Tracy, & Singer, 1985). Long rank list of different crimes by seriousness have been generated, and index values (the seriousness scores) resulting have in fact been put into use in some jurisdictions as a tool of crime measurement or trend quantification (Burton et al., 2004). Methodological studies on the other hand have explored different methods of developing crime seriousness ranking or scores (index values), in an attempt to develop a more robust tool for this type of research.

It should be noted that the criminological literature appears to have taken the concept of crime seriousness at its face value, and there has been little focus on deconstructions of the construct itself (Stylianou, 2003; Walker, 1978). For example, Rossi, Waite, Bose and Berk (1974, p.231) admit to neither specifying what they mean by seriousness, nor asking study participants what they understand by the term *seriousness*. Evans and Scott (1984) asked participants to assign penalties, which were assessed by their severity, to offences, thereby defining offence seriousness through punishment perceived as deserved. A similar approach was used by Skovron, Scott and Rao (1987), in a comparative study of crime seriousness in the United States, India and Kuwait.

One of the first more structured attempts to understand what seriousness assessments entails was by Fishman et al. (1986). In their study they had found that personal injury and criminal intent were the two components of seriousness assessments, accounting for 78% of variance in seriousness scores. They had also established that these two dimensions were independent (not correlated), while both were separately positively

correlated with seriousness, further supporting their argument that seriousness is a two-dimensional construct.

Warr (1989) proposed that judgments of crime seriousness are a combination of harmfulness and wrongfulness judgements. His findings, however, suggested that the degree to which harmfulness and wrongfulness contributed to the perceived crime seriousness varied by types of offence. In particular, he found that for violent offences, harmfulness was a stronger predictor of perceived seriousness of the crime, while for property offences, wrongfulness was the stronger predictor. O'Connell and Whelan (1996) have discussed seriousness assessments in terms of "badness" and "individual impact" dimensions. They had found that "badness" accounted for 42,3% while "individual impact" for 10,4% of the variance in seriousness assessments.

These research findings indicate that perceptions of crime seriousness are framed by what can be summarized as two dimensions: The consequences of the act, and the moral wrongfulness of the act, which also is the basic premise of Warr's model. In other words, Warr conceptualized crime seriousness as a reflection of both normative evaluations regarding moral gravity of the offence, and consequential characteristics of a crime (factual assessment of consequences for the victim).

Empirical tests of this two-dimensional model have provided considerable empirical support (O'Connell & Wheelan, 1996; Rosenmerkel, 2001; Alter, Kernochan, & Darley, 2007; Fishman, Kraus & Cohen, 1986; Adriaenssen, Paoli, Karstedt, Visschers, Greenfield, & Pleysier, 2018), allowing it to become the dominant way of conceptualizing crime seriousness. And yet even the original study by Warr had reported that for some groups, and for some offences, the two-dimensional model may not be a good fit. For example, religiosity was found to be associated with less differentiation in wrongfulness assessments, meaning that for those for whom moral values are strongly set by religion crime seriousness may be a one-dimensional concept, defined primarily by harmfulness (Warr, 1989). Findings also suggest that for some offences (such as common street crimes) seriousness is closely associated with wrongfulness, while for others (such as white collar offenses) it is closely associated with harmfulness (Rosenmerkel, 2001), suggesting that the construct of seriousness may rely on different set of evaluations for different types of offenses. In fact, a number of studies (Adriaenssen et al., 2018; Alter et al., 2007; O'Connell & Wheelan, 1996) have found wrongfulness to be a stronger predictor of offence seriousness, concluding that consequentialism is the less dominant approach to assessing the seriousness of crime.

When it comes to the degree to which there is an agreement among individuals and groups regarding how serious different offences are, studies indicate that there is a rather strong consensus regarding seriousness of offences both within as well as across cultures (Stylianou, 2003). Whether the same level of agreement can be found for judgments of harmfulness and wrongfulness, and their relative contribution to the perceived seriousness, has not been tested as extensively in a cross-cultural context. So far, studies testing Warr's model were conducted in the USA (Warr, 1989; Rosenmerkel, 2001), Ireland (O'Connell & Whelan, 1996), and Belgium (Adrianssen, Paoli, Karstedt, Visschers, Greenfield, & Pleysier, 2018), and the common result in all these studies is that overall perceived wrongfulness is a stronger predictor of the offence seriousness than harmfulness is. In a study that explored the relationship between harmfulness and wrongfulness of an offense and proposed sentence severity (Alter, Kernochan, & Darley, 2007) it was similarly found that wrongfulness predicted severity of a "deserved" sentence better than harmfulness. However, results also suggest that dominance of wrongfulness as predictor of perceived crime seriousness is not absolute and varies by the type of offence. Thus for example, Warr found wrongfulness to be the sole predictor of crime seriousness for property related offences, while harmfulness was twice as strong a predictor as wrongfulness for offences that are violent in nature or present a risk to life (1989, p. 809), and Rosenmerkel found that for white collar crimes harmfulness was a stronger predictor of seriousness assessments (2001, p. 322). In study by Andrianssen et al. (2018), while for different offences the predictive strength of wrongfulness and harmfulness was different from offence to offence, for none of the offences was harmfulness a stronger predictor than wrongfulness (2018, p. 16-17).

The purpose of this study is to contribute to the discussion of crime seriousness from a cross-cultural viewpoint, by studying the issue with a Turkish sample. Consensus regarding crime seriousness has already been established in a cross-cultural setting (Stylianou, 2003), including in majority Muslim societies (Skovros, Scott, & Rao, 1987), which suggests that there is a general agreement regarding which crimes are perceived as more and which as less serious across different societies and cultures. Evans and Scott (1984), however, also found that what they labeled as "moral offences" to be perceived as more serious in a Kuwait sample compared to the sample from the USA. This outcome, however, was concluded to be the function of religious fundamentalism, rather than cultural values or the religion per se, thus not so much cultural difference, as a matter on individual variability related to the degree of religiousness. While Curry (1996) had found that conservative religious beliefs were

associated with less discrimination regarding wrongfulness, Adriaenssen et al. (2019) found that conservatism was associated with perceiving offences as more serious, while the similar effect was not as consistent for religiosity. These findings suggest that examinations of the Warr's model in different cultural settings might lead to findings that suggests that the model is not in fact universal, but is rather shaped by religious beliefs, degree of conservatism of the population, and other cultural and social variables which could potentially vary from country to country, culture to culture. So far Warr's model of offence seriousness has been tested only in a small number of cultures (USA, Ireland, and Belgium-Flanders) that were not that dissimilar from one another, and not at all in a majority Muslim society yet. Hence the goal of this study is to provide results from yet another culture regarding the degree of consensus regarding perceptions of the seriousness of crime, and to test Warr's model of crime seriousness in a different cultural setting. In particular, the consensus regarding crime seriousness, and wrongfulness and harmfulness perceptions were studied, as well as whether the wrongfulness is the stronger predictor of crime seriousness assessments, as has been found in studies conducted so far.

2. Method

Items used in the Warr's 1989 study were translated to Turkish by two translators independently. Translations were then compared, and problematic items were reviewed and consolidated. The final version of the translation of the items is in Annex I.

Students registered at a privately-owned university in Istanbul were informed via BlackBoard announcement page of a possibility to participate in the study, during the Spring semester of 2019. Participation was voluntary, and while some were offered course credit in exchange for participation, some were not. In order to participate students needed to click on a link, that led them to an on-line questionnaire, at their own time.

In the on-line questionnaire, a short introduction to the study and the consent form were followed by questions related to age, gender, field and year of study. Following these questions, the same procedure was followed for the three stages of the survey, each asking participants to assess listed offences based on a different criterion: seriousness, wrongfulness, and finally harmfulness.

First, the participants were provided with an explanation of the task. Instruction were very similar to those used in Warr's study but adapted for on-line modality of the questionnaire. The instruction text can be found in Annex II. Following this instruction,

on the next page a multiple-choice question was asking the participant to confirm the criteria that they are asked to use when assessing the offences (this question is presented in Annex III). If the answer provided was wrong, a page popped up informing the participant that their answer was wrong, and restating again the instruction (these additional instructions can be seen in Annex IV).[2] This was followed by a page that listed all offences at the same time, and the participants were able to click on a value 0-10 for each offence. While the order of items was the same as that used in the Warr study, participants were able to see all offences at the same time, and choose values in any order they want, and also change their choices, before submitting their answers. This procedure was followed for each of the three assessment criteria, and the same order of assessment was used for all participants: First seriousness, followed by wrongfulness, and finally harmfulness. Data obtained were downloaded and analyzed using SPSS software.

Original items 7, 8, 14, 15, 17, 20, 22, and 28 refer to an identity of the offender (a teenager, a father, a parent, etc.), while other items do not. Translation into Turkish makes this difference in item structure even more apparent, as grammatical structure of those items is different. In order to test whether providing *some* information regarding the identity of the offender has an influence on the assessments of the offences, an alternative version of the questionnaire was developed as well, in which offender characteristics were removed all together from the items. Those alternative versions can be seen in the Table 1. Each participant only got one version of the items.

Table 1. Alternative version of the items, in English and in Turkish

Original Item No	Adapted offence description	Turkish translation
7	Hitting an old woman in the street	Sokakta yaşlı bir kadına vurmak
8	Sexually abusing own teenage daughter	Kızını cinsel olarak istismar etmek
14	Overcharging $60 on auto repairs	Oto tamiri faturasını 600 TL değerinde şişirmek
15	Beating own child with fists	Çocuğunu yumruklayarak dövmek
17	Beating up a classmate in highschool	Lisede sınıf arkadaşını dövmek
20	Running hands over a woman on a bus	Kadını otobüste ellerle taciz etmek
22	Hitting one's wife during an argument	Tartışma sırasında karına vurmak
28	Forcing one's girlfriend into sex	Kız arkadaşını cinsel ilişkiye zorlamak

---

2   Less than 6% of participants answered any of those three questions wrong, indicating that participants were paying attention to the instructions and were aware which criteria they were expected to use in each step when assessing offences.

3. Participants

149 participants took the original survey, while 161 participant took the survey with alternative items, totaling 310 participants. It took them on average 12 minutes to complete the survey. A total of 258 participants provided answers to all the questions (items) in the survey. Average age was 22,03 (SD=3,58), 68,7% were female, and 67,3% of students were in law programs (undergraduate or graduate). Participants were not asked to provide information regarding their religious beliefs. However, according to findings from World Values Survey, 99% of individuals in Turkey self-identify as Muslim (question v144) and 83,5% identify as "religious person" (question v147) (Inglehart, Haerpfer, Moreno, Welzel, Kizilova, Diez-Medrano, Lagos, Norris, Ponarin & Puranen et al., 2014), meaning that it is safe to assume that the sample in this study was predominantly Muslim.

4. Results

4.1. Assessing Alternative Versions of Items

In order to test whether the wordings used in the original and the alternative versions of the items resulted in differences in assessments, independent sample T-test was applied to individual items, for all three assessment criteria separately. When seriousness assessments were compared, significant difference between means was found only for item 17 (original item X = 6,91; SD = 2,77; alternative item X = 4,58; SD = 2,86; t (308) = -7,89, p < 0,000). For wrongfulness assessments, the results were similar, with item 17 being the only one with statistical significant difference between the means of the two versions of the item (original item X = 8,13; SD = 2,24; alternative item X = 6,37; SD = 3,32; t (275) = -5,06, p < 0,000). For harmfulness assessments, in addition to item 17 (original item X = 7,98; SD = 2,44; alternative item X = 5,96; SD = 3,18; t (256) = -5,69, p<0,000), a statistically significant difference between the means of the two versions of the item was found for the item 28 as well (original item X = 8,84; SD = 1,94; alternative item X = 9,26; SD = 1,44; t (256) = 2,00, p < 0,05). Due to the high number of the T tests performed, it was assessed that the difference between means related to harmfulness assessments for the item 28 was likely due to chance. Since the difference between assessment means for item 17 was consistent for all three types of assessment, this item was excluded from further analysis, while assessments for other items were combined into a single sample, which was used for further analysis.

4.2. Offence as Unit of Analysis

Figure 1 shows means seriousness assessments from this study for along with the means from Warr's 1989 study, for comparison purposes[3]. For most offences, mean seriousness assessments were similar, and the means in the two studies were highly correlated (Pearson's r = 0,94; p < 0,001). Items in which the difference between the means is more substantial are in marked in the chart.
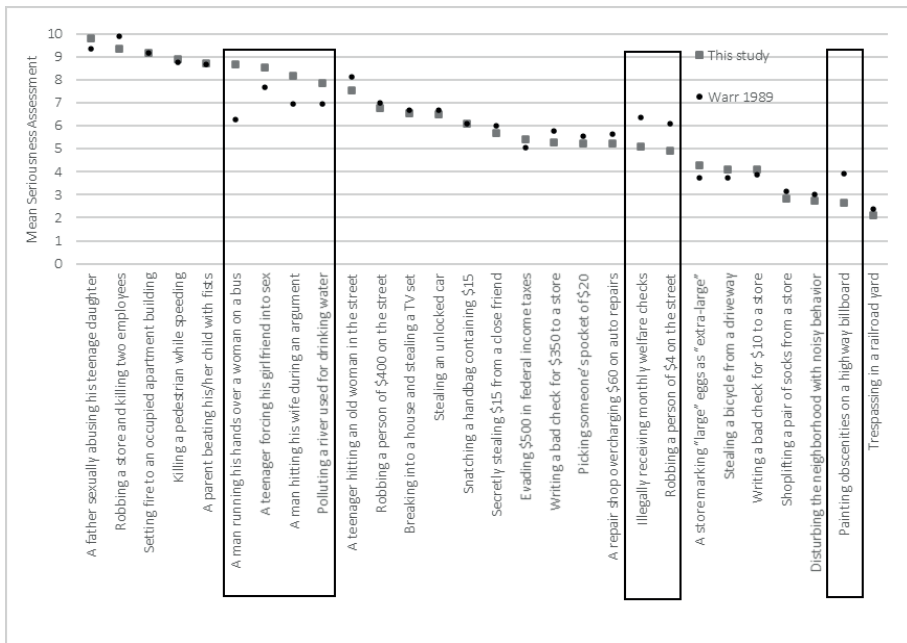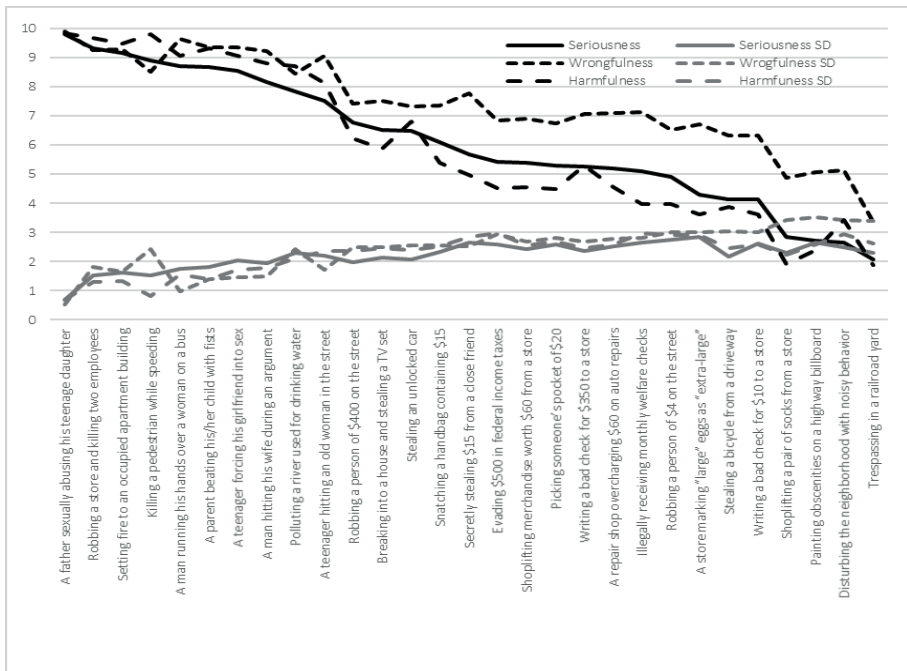

Figure 1. Mean seriousness assessments in this and Warr's (1989) study

Figure 2 shows means and standard deviations for seriousness, wrongfulness, and harmfulness assessments. Offences that involve violence and/or threat to life and bodily harm have been assessed as the most serious/wrongful/hurtful, followed by offences related to different types of theft and other property damage, and finally followed by what can be called public order offences and minor property related crimes. With the decrease in the seriousness assessment, the gap between the assessment of seriousness and wrongfulness increases.

---

3   Items for which monetary value was different in the Turkish version are not shown.

This chart demonstrates an inverse relationship between the means of perceived seriousness, wrongfulness, and harmfulness of offences, and the standard deviations, which are an indicator of agreement/consensus (Miethe, 1982). This suggests a high level of agreement when it comes to offences assessed as more serious, and much less agreement when it comes to offences that are assessed as less serious. As the perceived seriousness, harmfulness, or wrongfulness goes down, the level of consensus decreases, which is reflected in increased standard deviation values. In fact when means were correlated with standard deviations, for all three types of assessment the correlation was found to be statistically significant and negative (for seriousness Pearson's r = -0,77, p < 0,01; for wrongfulness Pearson's r = -0,92, p < 0,01; for harmfulness Pearson's r = -0,84; p < 0,01).

Figure 2. Means and standard deviations for seriousness, wrongfulness, and harmfulness assessments

4.3. Testing Warr's Model

Figure 3 shows harmfulness and wrongfulness means in a two-dimensional plot, similar to Warr's Figure 1 (1989, p. 803). While the regression equation is similar to that found by Warr (in this study WRG = 4,00 + 0,57 * HRM; in Warr's study it was WRG = 3,50 + 0,62 * HRM), clustering of the offences is different.
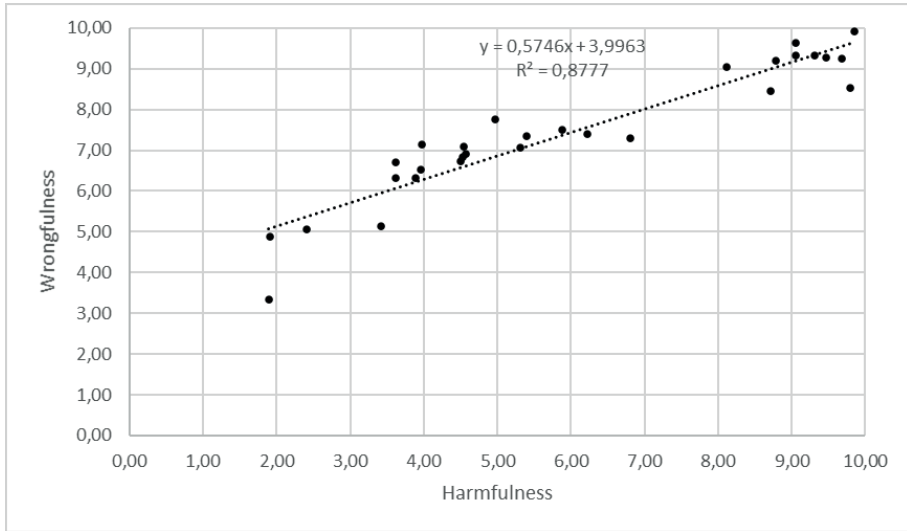
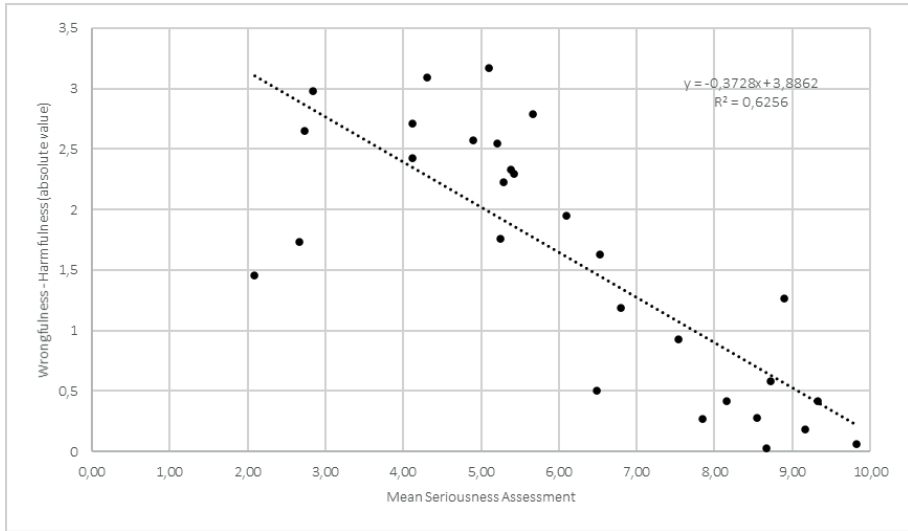**Figure 3. Harmfulness and wrongfulness assessment means**
(each dot represents different offence)

Figure 4 shows differences between wrongfulness and harmfulness mean assessments in this study and the Warr's study. As can be seen from this chart, in this study most offences were assessed as more wrongful than harmful, while in Warr's study there were more offences with higher harmfulness assessment, and the difference between harmfulness and wrongfulness assessments were overall smaller in Warr's study. Means of absolute value of differences between wrongfulness and harmfulness assessment in this study was 1,60, while in Warr's study it was 0,47, which suggest much larger difference between wrongfulness and meaningfulness assessment in this sample.
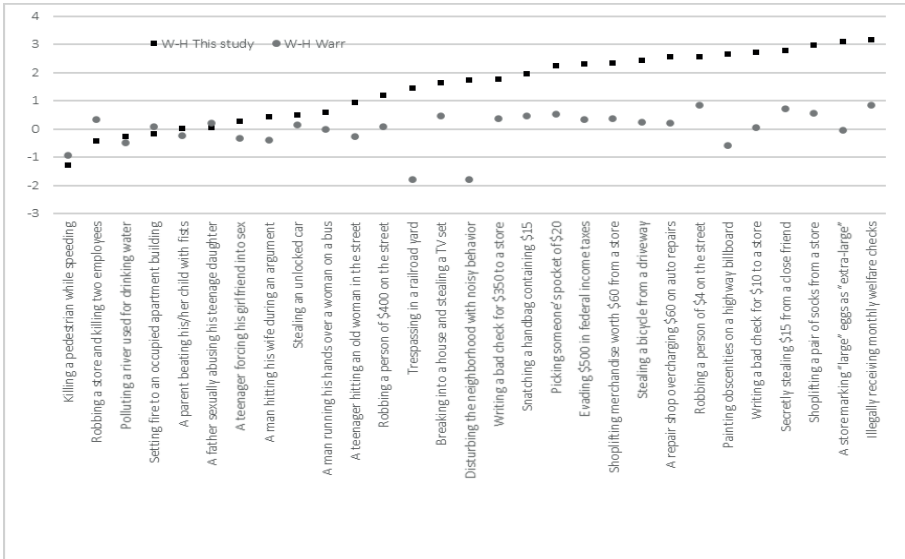
Figure 4. Difference between the means of wrongfulness and harmfulness assessments, in this study and the Warr's (1989) study (negative value indicates higher harmfulness than wrongfulness mean)

Differences between means of wrongfulness and harmfulness assessments found in this study (using absolute values) were plotted with mean seriousness assessments in the Figure 5. There is a clear negative relationship ($R^2 = 0,63$) between the two variables. This indicates that as assessments of seriousness increase, the agreement between harmfulness and wrongfulness increases as well, resulting in smaller difference between two assessments. In other words, offences perceived as more serious are more likely to be perceived as equally harmful and wrongful. The differentiation between the two types of assessment is more pronounced with offences that are seen as less serious.

Figure 5. Relationship between seriousness assessments and the difference between wrongfulness and harmfulness assessments (absolute values)

Clustering of variables in this study can be seen in the Figure 6, which shown results of hierarchical cluster analysis using Ward's method (1963) using wrongfulness and harmfulness.
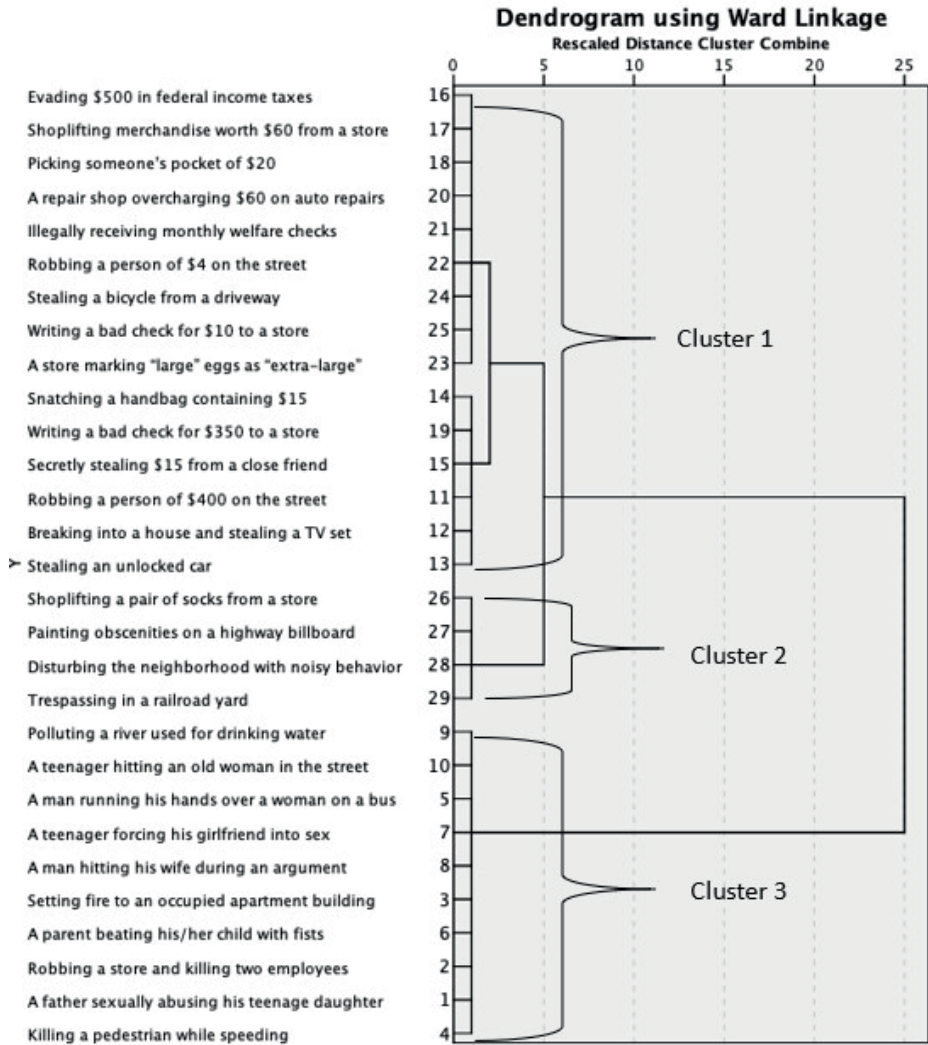
Figure 6. Results of hierarchical cluster analysis applied to harmfulness and wrongfulness means for offences, showing 4, 3 and 2 cluster solutions

When compared to clusters identified by Warr, while there is some overlap, categories are not identical. The main difference is that in this study there is no differentiation between what Warr had identified as Categories II and IV. Offences from those two Categories are all included into Cluster 3 here, which can be called "offences related to bodily injury/

death". Warr had separated those offences into Categories II and IV based on the difference between wrongfulness and harmfulness assessment, with Category II including those offences where the difference was positive (assessed as more wrong than harmful), and Category IV including those offences where the difference was negative (assessed as more harmful than wrong). In this study even though for some offences in Cluster 3 there was a negative difference between wrongfulness and harmfulness (meaning that they were assessed as more harmful than wrong), this did not result in a separate cluster.

Further, offences from Warr's Category I (property offences) are part of two sub-clusters in a four cluster solution (one including "More" and the other "Less" serious property offences). In a three-cluster solution these two clusters converge into a single Cluster 2 that almost fully overlaps with Warr's Category I.

Finally, Warr had named Category III "public offences". In this study, Cluster 2 includes all those offences, in addition to "shoplifting a pair of socks" item. This could mean this cluster has more to do with offences being perceived as minor, rather than as "public".

The two cluster solution separates variables into two basic categories: Those that involve threat/risk of physical harm/death, and the others.

Regression analysis (results shown in Table 2) applied to variables as cases suggests that wrongfulness and harmfulness do contribute equally to seriousness assessments, and harmfulness is somewhat stronger predictor of seriousness than wrongfulness. When the analysis is repeated for Cluster 3 offences only (threat/risk of physical harm/death offences), the difference is even more pronounced, and harmfulness is more than twice as strong predictor of seriousness as wrongfulness is. In Warr's study, harmfulness was also found to be a strong predictor of seriousness assessments for Category IV (which includes most, but not all offences in Cluster 3), but wrongfulness was not a significant predictor.

Table 2. Models predicting seriousness for different clusters of offences (unstandardized beta coefficients used in the formula)

Simple additive model for all offences	
$SER = -0{,}95 + 0{,}53 * WRG^{***} + 0{,}52 * HRM^{***}$ Standardized beta coefficients: Wrg 0,39 Hrm 0,62	Adjusted $R^2=0{,}97^{***}$
Simple additive model for Cluster 3 offences (bodily harm/life injury/risk) offences	
$SER = -5{,}98 + 0{,}52 * WRG^{***} + 1{,}07 * HRM^{**}$ Standardized beta coefficients: Wrg 0,34 Hrm 0,85	Adjusted $R^2=0{,}97^{***}$
Simple additive model for Cluster 1 (property) offences	
$SER = -1{,}88 + 0{,}63 * WRG^{*} + 0{,}59 * HRM^{***}$ Standardized beta coefficients: Wrg 0,32 Hrm 0,68	Adjusted $R^2=0{,}88^{***}$
*** p<0,001; ***p<0,01; * p<0,05	

For property offences, the model is somewhat weaker (even though still very strong), and harmfulness is a much stronger predictor (standardized beta) of seriousness assessments than wrongfulness. This is in stark contrast with Warr's results, where wrongfulness was found to be almost an exclusive predictor of seriousness assessments for property offences.

To summarize, in this study harmfulness and wrongfulness assessments were found to be strong predictors of seriousness assessments, regardless of the type of crime[4], yet harmfulness was a stronger predictor of seriousness than wrongfulness. This was especially pronounced with Cluster 3 (offences related to threat/risk of physical harm/death) offences.

4.4. Individual Variability

In the previous pages, results of analysis that used offence as a unit of analysis were shown. Here results of analysis for each offence separately will be presented as well, in order to determine whether findings that were found will also hold with individuals as the unit of analysis.

Strong correlation between seriousness and wrongfulness was found for all offences, with association ranging from r = 0,48 for "A man running his hands over a woman on a bus" to r = 0,71 for "Setting fire to occupied building". "Robbing a store and killing two employees" was an outlier with correlation of r = 0 ,28, which was still significant, but not as strong. This is likely due to very high level of agreement when it comes to this offence, meaning that there is little variance, and as a result it correlated not as strongly. Similarly, correlation between seriousness and harmfulness were all statistically significant as well and rather strong, ranging from r = 0,50 for "Stealing unlocked car" to r = 0,70 for "Writing a bad check". Outlier results of these analyses were "A father sexually abusing his teenage daughter" (r = 0,36), "Robbing a store and killing two employees" (r = 0,35), and "Killing a pedestrian while speeding" (r = 0,22), all of which were nevertheless statistically significant.

In order to test whether Warr's model holds on individual level as well, regression analyses were performed for each offence, with harmfulness and wrongfulness as independent variables, and seriousness assessment as a dependent variable. Those results are shown in Table 3. All models were significant (with p < 0,001) with adjusted R^2 values ranging from 0,16 ("Robbing a store and killing two employees") to 0,56

---

4   Due to a small number of offences/cases (only four), regression was not performed using Cluster 2 offences.

("Setting fire to an occupied building"), and with coefficients for both harmfulness and wrongfulness being highly significant for most offences. In order to explore the relative strength of harmfulness and wrongfulness as predictors of seriousness assessments, the difference between beta coefficients of the two types of assessments were plotted in Figure 7. This figure clearly illustrates that once again our findings differ from Warr's. In that study in only five offences harmfulness was stronger predictor of seriousness than wrongfulness, and in three additional offences the beta coefficients were equal. In this study there is an even split, with harmfulness being stronger predictor than wrongfulness in 14 offences (offences on the right). This is a clear indicator that in population that this sample was derived from harmfulness plays a more prominent role in assessments of seriousness, and is not a secondary predictor.

Table 3. Results of regression of Seriousness on Harmfulness (HRM) and Wrongfulness (WRG) for each offense

Offence	Adj. R²	WRG Beta	HRM Beta
Robbing a store and killing two employees	0,16	0,21	0,30
A father sexually abusing his teenage daughter	0,25	0,43	0,11†
Stealing a bicycle from a driveway	0,35	0,37	0,32
Stealing an unlocked car	0,37	0,42	0,26
Breaking into a house and stealing a TV set	0,38	0,41	0,30
Killing a pedestrian while speeding	0,38	0,59	0,12*
Writing a bad check for $20 to a store	0,38	0,37	0,35
Secretly stealing $15 from a close friend	0,38	0,31	0,41
A teenager hitting an old woman in the street	0,38	0,36	0,35
Shoplifting a pair of socks from a store	0,39	0,44	0,28
Robbing a person of $400 on the street	0,41	0,29	0,43
Painting obscenities on a highway billboard	0,42	0,40	0,37
Robbing a person of $4 on the street	0,43	0,31	0,43
Illegally receiving monthly welfare checks	0,44	0,42	0,36
Snatching a handbag containing $15	0,44	0,28	0,47
A parent beating his/her child with fists	0,45	0,57	0,15**
Trespassing in a railroad yard	0,45	0,36	0,39
A man running his hands over a woman on a bus	0,45	0,20	0,54
A man hitting his wife during an argument	0,45	0,36	0,38
Evading $500 in federal income taxes	0,46	0,44	0,33
A repair shop overcharging $60 on auto repairs	0,48	0,31	0,46
Polluting a river used for drinking water	0,49	0,43	0,36
Picking someone's pocket of $20	0,50	0,39	0,43
Disturbing the neighborhood with noisy behavior	0,50	0,30	0,49
A teenager forcing his girlfriend into sex	0,50	0,32	0,43
Shoplifting merchandise worth $60 from a store	0,51	0,34	0,47
Writing a bad check for $350 to a store	0,53	0,27	0,53

A store marking "large" eggs as "extra-large"	0,55	0,37	0,49
Setting fire to an occupied apartment building	0,56	0,58	0,27
All models are significant with p<0,001; All coefficients are standardized, significant at p<0,001, unless where noted otherwise; † p>0,05; * p<0,05; ** p<0,01			
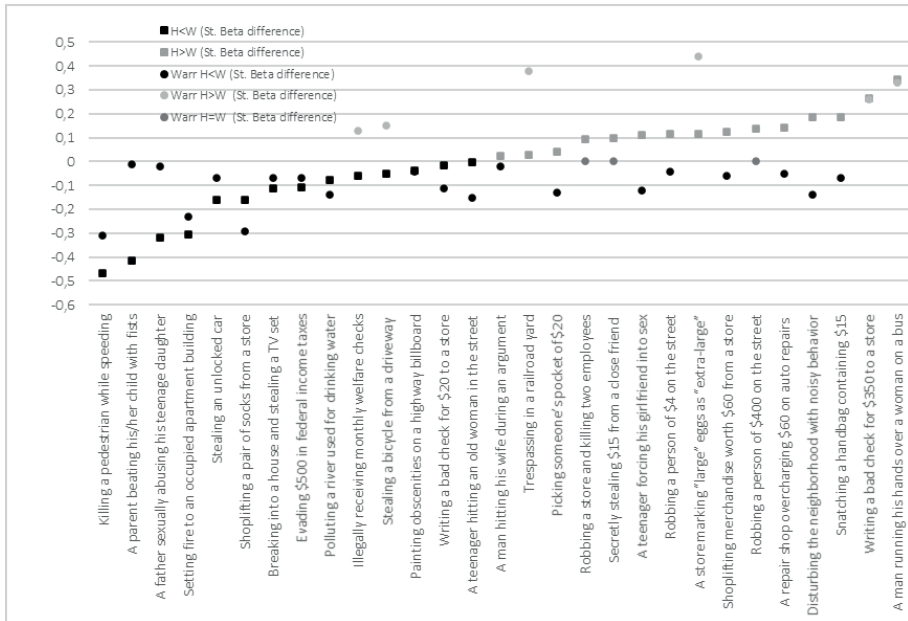
Figure 7. Difference between standard beta coefficients for harmfulness and wrongfulness in this study and Warr's study

For clarity, main research questions and relevant findings are summarized in Table 4.

Table 4. Main research questions and summary of findings

How seriousness assessments in this study compare with those from Warr's study?	High correlation between the two studies. Seriousness assessment means similar for most offences, exceptions noted.
What is the level of consensus within the sample regarding seriousness, wrongfulness and harmfulness of offences?	High level of consensus for offences assessed as serious; much lower (high SDs) for offences assessed as less serious (significant negative correlation between means and SDs)

How harmfulness and wrongfulness assessments in this study compare with those from Warr's study?	In this study, most offences assessed as more wrong than harmful, with larger difference between two assessment. Higher level of similarity between seriousness and wrongfulness assessments for offences judged as more serious.
How did offences cluster?	Three main clusters emerged: violent/physical harm offences, property crimes (with two subclusters: "more" and "less" serious), and minor offences.
Do wrongfulness and harmfulness predict serious-ness assessment (does Warr's model hold)?	Yes. However, harmfulness much stronger predictor of seriousness than wrongfulness for both violent and property offences (in contrast to findings from other countries).
On individual level, does two-dimensional model (Warr's model) hold?	Yes, for most offences. Still, harmfulness found to be the prominent (rather than secondary) predictor (in contrast to findings from other countries)

5. Discussion and Conclusion

By assessing alternative version of items, it was found that minor changes in language did not make much difference when it comes to how offences are assessed. The questions of how such information influences assessment has been discused before. Wolfgang et al. (1985) had found that offences that involved identical criminal act (stabbing) and identical outcomes (death) were rated differently depending on the gender of the offender and the victim. However, the case in which the gender was not known was rated more similarly to the one with male offender and female victim, suggesting that even when information about the offender and the victim is not provided the assessments are not made in a vacuum. Lynch and Danner (1993) had found that a number of characteristics of the event, including the gender of the victim and the offender, as well their age, have an effect on the seriousness ratings. When such information is not provided, people base their assessments regarding seriousness of offences on stereotypical representation of these offences that they have (Parton, Hansel, & Stratton, 1991), which means that such assessments are never really made in an abstract. In other words, even if the item stated "a person is sexually harassed on a bus", it will likely be assessed as "a woman is sexually harassed on a bus by a man", based on ideas that people have about how such offences happen in reality, true or not. Hence, whether it is specified that "a man" has sexually harassed a woman or not becomes irrelevant, as the crime of "sexual harassment" will likely be assumed to be committed by a man against a woman, and assessed as such. Consequently, the presence or the lack of specific information regarding the victim or offender's gender

may not make much difference, unless if the information provided is in contrast with the stereotype ("a woman sexually harassed a man on the bus").

Similar argument can be made for most alternative items that were tested in this study. The fact that the "offender" in the offence was not clearly defined does not mean that participant did not derive who that offender could be (a parent, a man, etc.). This does not mean however that some other alternative versions would not illicit a different response. For example, if the item regarding hitting an old woman specified gender or age of the offender, that could potentially lead to different assessments. In this particular research we did not test whether an offender/victim different from stereotypical one would lead to different assessments, and thus this remains to be explored in future research.

Comparison between results of this study and Warr's with regard to how different offences are assessed in terms of their seriousness once again confirms that there is a high degree of consensus regarding crime seriousness between cultures, for most offences, with some exceptions noted. Where there were differences, those could be attributed to a number of factors: cultural differences, change over time, sampling or measurement error. In this particular study, seriousness of sexual harassment, sexual assault, and domestic violence were all assessed as more serious than what was the case in Warr's study. Collectively these three offences all entail violence against women. It would be optimistic to argue that among the Turkish population violence against women is perceived in a more negative way, and that assessments are due to higher levels of sensitivity regarding such violence in Turkey. While violence against women has been a topic of heated discussion over the last decade, and awareness has certainly been raised, higher seriousness assessments could just as likely be due to the characteristics of this particular sample. In this study sample was drawn from young, high SES, urban, and educated population, which does not represent the Turkish population in general, or even a university student population in particular. Variables such as age (O'Connell & Wheelan, 1996) and poverty (Levi & Jones, 1985) have already been demonstrated to be associated with the judgments of seriousness. Further, it must be noted that over the last 30 years, awareness regarding violence against women and sensitivity towards it has improved globally, and if the Warr's study was repeated today with similar sample in the USA, it would not be surprising if the offences against women were rated as more serious than in 1989.

However there are also basis to entertain the possibility that the difference is actually culture based. Evans and Scott (1984) in their research had found that sexual offences

(such as forcible rape and forced prostitution) were assessed as more serious in Kuwait, compared to the USA (p. 48). Given that Kuwait is a predominantly Muslim country, this could be relevant for our findings as well. Perhaps this points to a religion-based sensitivity to such offences. But the difference between the two countries was more pronounced with acts such as prostitution, adultery, and homosexuality, suggesting that the difference in the judgments of seriousness between the USA and Kuwait samples in relation to rape and forced prostitution was more likely to be associated with perceived immorality related to the *sexual* dimension of those offences, rather than the *violence against the women* dimension. This explanation is further supported by the findings that in non-sexual offences that entailed violence against women (such as killing or stabbing a wife), there was no difference between the two samples. Data from the World Values Survey 2010-2014 suggest that, in fact, Turkey is more similar to the USA, then to Kuwait, when it comes to attitudes towards domestic violence, with the mean assessment from the USA and Turkey being very similar and suggesting more negative attitudes than those found in Kuwait (Inglehart, Haerpfer, Moreno, Welzel, Kizilova, Diez-Medrano, Lagos, Norris, Ponarin & Puranen et al., 2014). Nevertheless, this remains a point that needs to be explored in the future, with both a more nationally representative sample in Turkey, and with more contemporary data from other countries with different religious and cultural backgrounds.

When it comes to within sample consensus, levels of consensus regarding seriousness, harmfulness, and wrongfulness of offences were not identical across all types of offences. Degree of consensus in offences that were assessed as very serious/harmful/wrongful was very high, and very low in offences at the opposite end of the spectrum. In other words, assessments and degree of agreement are inversely related. Similarly, for offences that were assessed as more serious, assessments of harmfulness and wrongfulness assessment were more similar as well (means were similar). In other words, not only was there more agreement about how serious/harmful/wrongful these offences are, they were also assessed as more similar in terms of their wrongfulness and harmfulness. It is possible to speculate that for serious offences, there is little distinction of severity between harmfulness and wrongfulness. Such offences (in this category were ones that involved violence, physical harm, risk to life and similar) are perhaps simply seen as "awful", using a sort of moral intuition (Haidt, 2001) instead of engaging into separate assessments of wrong and harm of the act. One could possibly argue that only offences that are seen as both very harmful and very wrong are therefore perceived as serious. However, Warr's two-dimensional model does not require this

unison of assessment. Theoretically, one could have an offence that is perceived as very wrong but less harmful, and still be assessed as rather serious (for example assed as 10 on wrongfulness, 6 on harmfulness, and 8 on seriousness). In this study, however, there were no offences at the "very serious" end of the spectrum with that kind of difference between the harmfulness and the wrongfulness assessment (Figure 4). In other words, serious offences are perceived as both wrong and harmful, and never as dominantly harmful or dominantly wrongful. Manipulating level of wrongfulness (through manipulation of intent) while keeping the harm identical would be an interesting way to explore whether there can be very serious crimes that are perceived as less wrong while equally harmful. It is already established (Sebba, 1980) that presence of intent leads to higher crime seriousness assessments. How this influences assessments of wrongfulness and harmfulness of an offence has not been researched yet.

With property related offences, there was more difference between harmfulness and wrongfulness assessments (Figure 2), and harmfulness assessment were more similar to seriousness assessments than wrongfulness assessments were. In other words, there was more differentiation between wrongfulness and harmfulness in offences that are perceived as less serious, and perhaps this is why they are seen as such.

These findings were also reflected in the results of the regression analysis with offence as the unit of analysis. First to reiterate that Warr's model has been supported by the findings in this study, both on the offence as well as on the individual level of analysis, and both harmfulness and wrongfulness were found to be reliable predictors of seriousness assessment for almost all offences. However, in this study the two predictors were not equal and harmfulness was a stronger predictor of seriousness regardless of the type of offence. The difference between the predictive power of the two dimensions was biggest with Cluster 3 offences, which are also the most serious offences. These findings are in stark contrast with the findings of recent research conducted in Belgium, where wrongfulness was a stronger predictor of seriousness for all types of offences (Adriaenssen, Paoli, Karstedt, Vissechers, Greenfield, & Peysier, 2018). At this point it is hard to assess what is behind this clear difference in dominating component between the two countries. Once again, it is highly likely that data obtained from a national sample would paint a different picture, and the findings reported here are characteristic to young, urban, educated population, that is perhaps more in tune with the concept of victimization and social harm caused by the crime. But it could also be an issue of differing cultural values.

While this study provides further support for Warr's model of components of crime seriousness, the questions that it raises are significant. While harmfulness and wrongfulness are good predictors or seriousness, exploring what on individual level predicts them is the next step that will need to be taken. This is closely related to the question of consensus. Among offenses assessed as very serious, the consensus levels were high, and it is unlikely that any demographic or attitudinal variables would have make much difference. On the other hand, standard deviations of seriousness assessment of offences assessed as less serious were so high that they beg for an explanation: What shapes perceptions of seriousness of offences that are not simply "awful"? Such high variability in seriousness assessments among individuals suggests that individual level explanations need to be employed.

Higher conservation values and religiosity were found to be positively associated with judgments of seriousness (Adrianssen, Karsted, Paoli, & Visschers, 2019), and it was argued that both are a reflection of valuing order. Age, gender, and socioeconomic status were found to be somewhat unreliable predictors of seriousness, harmfulness and wrongfulness assessment, but for some offences they did produce an effect. Based on these findings there are two clear line of research that need to be pursued in order to further contribute to our understanding regarding what shapes perception of crime severity: first one is in relation to characteristics of the offences, and the second one is in relation to individual differences. First, it is important to try to study separately the effect of harmfulness and wrongfulness on seriousness perceptions, by manipulating them separately and examining how this manipulation influences seriousness assessments. This will truly show whether these two dimensions are independent of one another, and perhaps demonstrate that for some crimes they are simply seen as one and the same and people do not really engage in any higher-level reasoning when making those assessments. The second line of research would need to focus on identifying individual characteristics that influence how people perceive crimes, by focusing on both attitudinal as well as demographic characteristics, in order to improve our understanding regarding what lies behind large individual difference in offences perceived as somewhat less serious. All this must be done for different categories of offences separately, as both previous as well as our research findings appear to suggest that different categories of crime are not assessed using the same mental calculus. Finally, the study must be expanded into other countries, in order to obtain more data points that will allow to better cross-cultural comparisons.

References/Kaynakça

Adriaenssen, A., Karstedt, S., Paoli, L., & Visschers, J. (2019). Taking Crime Seriously: Conservation Values and Legal Cynicism as Predictors of Public Perceptions of the Seriousness of Crime. *International Criminal Justice Review*, *29*(4), 317-334.

Adriaenssen, A., Paoli, L., Karstedt, S., Visschers, J., Greenfield, V. A., & Pleysier, S. (2018). Public perceptions of the seriousness of crime: Weighing the harm and the wrong. European Journal of Criminology, 17(2), 127-150. https://doi.org/10.1177%2F1477370818772768.

Alter, A. L., Kernochan, J., & Darley, J. M. (2007). Transgression wrongfulness outweighs its harmfulness as a determinant of sentence severity. *Law and Human Behavior*, *31*(4), 319-335. https://doi/10.1007/s10979-006-9060-x.

Burton, S. E., Finn, M., Livingston, D., Scully, K., Bales, W. D., & Padgett, K. (2004). Applying a crime seriousness scale to measure changes in the severity of offenses by individuals arrested in Florida. *Justice Research and Policy*, *6*(1), 1-18. https://doi.org/10.3818%2FJRP.6.1.2004.1.

Carlson, J. M., & Williams, T. (1993). Perspectives on the Seriousness of Crimes. *Social Science Research*, *22*(2), 190-207. https://psycnet.apa.org/doi/10.1006/ssre.1993.1009.

Cullen, F. T., Link, B. G., & Polanzi, C. W. (1982). The Seriousness of Crime Revisited: Have Attitudes Toward White-Collar Crime Changed?. *Criminology*, *20*(1), 83-102. https://psycnet.apa.org/doi/10.1111/j.1745-9125.1982.tb00449.x.

Cullen, F. T., Link, B. G., Travis, L. F., & Wozniak, J. F. (1985). Consensus in crime seriousness: Empirical reality or methodological artifact?. *Criminology, 23*(1), 99-117. https://doi.org/10.1111/j.1745-9125.1985.tb00328.x.

Curry, T. R. (1996). Conservative Protestantism and the perceived wrongfulness of crimes: A research note. *Criminology, 34*(3), 453-464.

Evans, S. S., & Scott, J. E. (1984). The seriousness of crime cross-culturally: The impact of religiosity. *Criminology*, *22*(1), 39-59. https://psycnet.apa.org/doi/10.1111/j.1745-9125.1984.tb00287.x .

Fishman, G., Kraus, V., & Cohen, B. Z. (1986). A multidimensional approach to the problem of crime seriousness. *International Journal of Comparative and Applied Criminal Justice*, *10*(1-2), 177-191. https://doi.org/10.1080/01924036.1986.9688842 .

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Herzog, S., & Einat, T. (2016). Moral judgment, crime seriousness, and the relations between them: An exploratory study. *Crime & Delinquency*, *62*(4), 470-500. https://doi.org/10.1177%2F0011128712466889 .

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin & B. Puranen et al. (eds.). 2014. World Values Survey: Round Six - Country-Pooled Datafile 2010-2014. Madrid: JD Systems Institute. Version:http://www.worldvaluessurvey.org/WVSDocumentationWV

Kwan, Y. K., Ip, W. C., & Kwan, P. (2000). A crime index with Thurstone's scaling of crime severity. *Journal of Criminal Justice*, *28*(3), 237-244. http://dx.doi.org/10.1016%2FS0047-2352%2800%2900039-8.

Lesieur, H. R., & Lehman, P. M. (1975). Remeasuring delinquency: A replication and critique. *The British Journal of Criminology*, *15*(1), 69-80. https://doi.org/10.1093/oxfordjournals.bjc.a046611 .

Levi, M., & Jones, S. (1985). Public and police perceptions of crime seriousness in England and Wales. *The British Journal of Criminology*, *25*(3), 234-250. https://doi.org/10.1093/oxfordjournals.bjc.a047529 .

Miethe, T. D. (1982). Public Consensus on Crime Seriousness Normative Structure or Methodological Artifact?. *Criminology*, *20*(3/4), 515-526. https://doi.org/10.1111/j.1745-9125.1982.tb00475.x .

O'Conell, M., & Whelan, A. (1996). Taking wrongs seriously: Public perceptions of crime seriousness. *The British Journal of Criminology*, *36*(2), 299-318. https://doi.org/10.1093/oxfordjournals.bjc.a014087 .

Parton, D. A., Hansel, M., & Stratton, J. R. (1991). Measuring crime seriousness: Lessons from the national survey of crime severity. *The British Journal of Criminology*, *31*(1), 72-85. https://doi.org/10.1093/oxfordjournals.bjc.a048086 .

Rose, A. M., & Prell, A. E. (1955). Does the punishment fit the crime? A study in social valuation. *American Journal of Sociology*, *61*(3), 247-259. https://psycnet.apa.org/doi/10.1086/221736 .

Rosenmerkel, S. P. (2001). Wrongfulness and harmfulness as components of seriousness of white-collar offenses. *Journal of Contemporary Criminal Justice*, *17*(4), 308-327. https://psycnet.apa.org/doi/10.1177/1043986201017004002 .

Rossi, P. H., Waite, E., Bose, C. E., & Berk, R. E. (1974). The seriousness of crimes: Normative structure and individual differences. *American Sociological Review*, 224-237. https://psycnet.apa.org/doi/10.2307/2094234

Skovron, S. E., Scott, J. E., & Rao, P. K. (1987). Cross Cultural Perceptions of Offense Severity: The United States, India and Kuwait. *International Journal of Comparative and Applied Criminal Justice*, *11*(1-2), 47-60. https://doi.org/10.1080/01924036.1987.9688855 .

Sebba, L. (1980). Is Mens Rea a component of perceived offense seriousness?. *The Journal of Criminal Law and Criminology*, 71(2), 124-135. https://doi.org/10.2307/1142721 .

Sellin, T., & Wolfgang, M. E. (1964). *The Measurement of Delinquency*. New York: Wiley.

Stylianou, S. (2003). Measuring crime seriousness perceptions: What have we learned and what else do we want to know. *Journal of Criminal Justice*, *31*(1), 37-56. https://psycnet.apa.org/doi/10.1016/S0047-2352(02)00198-8

Taneri, G. (2016). Temel cezanın belirlenmesi. *Ankara Barosu Dergisi, 3*, 127-161.

Üzülmez, İ (2006). Yeni Ceza Kanunu'nun Sisteminde Cezanın Belirlenmesi ve Bireyselleştirilmesi. *Erzincan Üniversitesi Hukuk Fakültesi Dergisi*, *10*(3-4), 203-235.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*, 236–244

Walker, M. A. (1978). Measuring the seriousness of crimes. *The British Journal of Criminology, 18*(4), 348-364. https://doi.org/10.1093/oxfordjournals.bjc.a046933 .

Warr, M. (1989). What is the perceived seriousness of crimes?. *Criminology*, *27*(4), 795-822. https://doi.org/10.1111/j.1745-9125.1989.tb01055.x .

Wolfgang, M., Figlio, R. M., Tracy, P. E., & Singer, S. J. (1985). The National Survey of Crime Severity (NCJ-96017). Washington, DC: US Government Printing Office. https://www.ncjrs.gov/pdffiles1/Digitization/96017NCJRS.pdf

Annex

Annex I. Original item description, and translations to Turkish that were used in this study

Original item no	Offence description (original item)	Item no in this study	Turkish translation
1	Robbing a store and killing two employees	1	Dükkan soyup, iki çalışanı öldürmek
2	Stealing a bicycle from a driveway	2	Evin önünden bisiklet çalmak
3	Shoplifting a pair of socks from a store	3	Mağazadan bir çift çorap çalmak
4	Robbing a person of $400 on the street	4	Sokakta birinden 4.000 TL gasp etmek[5].
5	Polluting a river used for drinking water	5	İçme suyu için kullanılan bir nehri kirletmek
6	Breaking into a house and stealing a TV set	6	Zorla birinin evine girip televizyon setini çalmak
7	A teenager hitting an old woman in the street	7	Ergenlik çağındaki gencin sokakta yaşlı bir kadına vurması
8	A father sexually abusing his teenage daughter	8	Bir babanın ergen yaştaki kızını cinsel olarak istismar etmesi
9	Stealing an unlocked car	9	Kilitlenmemiş bir arabayı çalmak
10	Evading $500 in federal income taxes	10	5.000 TL'lik gelir vergisi kaçırmak.

---

5   To adapt values of items stolen to Turkey while also taking into account purchasing power parity, we used Big Mac index (The Economist https://www.economist.com/news/2020/01/15/the-big-mac-index). In 1989 in USA, when original study by Warr was conducted, Big Mac was 2.02$. So for example, item 4 in which it is stated that 400$ was stolen, we calculated that the value of 400$ would be equivalent to the price of 200 Big Macs. In 2019 Big Mac in Turkey was priced at 18.99 TL, so amount of 3798 TL was equivalent to 400$ in 1989 in USA in terms of PPP. In order to round the amount, 4000 TL was used in the questionnaire. Same calculation was used for all items that include monetary value, and the amount was rounded up.

11	Killing a pedestrian while speeding	11	Hız yaparken bir yayayı öldürmek
12	Writing a bad check for $20 to a store	12	200 TL değerinde karşılıksız çek yazmak.
13	Secretly stealing $15 from a close friend	13	Yakın arkadaştan gizlice 300 TL çalmak
14	A repair shop overcharging $60 on auto repairs	14	Bir tamirhanenin faturayı 600 TL değerinde şişirmesi
15	A parent beating his/her child with fists	15	Bir ebeveynin çocuğunu yumruklayarak dövmesi
16	Setting fire to an occupied apartment building	16	Bir apartmanı, içinde insanlar varken kundaklamak
17	Ten high school boys beating up a classmate	17	On liseli erkeğin bir sınıf arkadaşlarını dövmesi
18	Robbing a person of $4 on the street	18	Sokakta birinden 40TL gasp etmek.
19	Illegally receiving monthly welfare checks	19	Kanuna aykırı bir şekilde sosyal yardım maaşı almak
20	A man running his hands over a woman on a bus	20	Bir adamın bir kadını otobüste elleriyle taciz etmesi
21	Snatching a handbag containing $15	21	İçinde 750 TL olan bir çantayı kapıp kaçmak
22	A man hitting his wife during an argument	22	Tartışma sırasında adamın eşine vurması
23	A store marking "large" eggs as "extra-large"	23	Marketin "büyük boy" yumurtaları "ekstra büyük" olarak pazarlaması

24	Disturbing the neighbor-hood with noisy behavior	24	Gürültü yaparak komşuları rahatsız etmek
25	Writing a bad check for $350 to a store	25	3500 TL değerinde karşılıksız çek yazmak
26	Trespassing in a railroad yard	26	Tren garına izinsiz olarak girmek
27	Picking someone's pocket of $20	27	Yankesicilik yaparak birinin cebinden 200 TL çalmak
28	A teenager forcing his girlfriend into sex	28	Ergenlik çağındaki gencin kız arkadaşını cinsel ilişkiye zor-laması
30[6]	Painting obscenities on a highway billboard	29	Otobanın kenarındaki reklam panosunun üstüne müstehcen şeyleri çizmek
31	Shoplifting merchandise worth $60 from a store[7]	30	Dükkandan 600 TL değerinde mal çalmak

Annex II. Instruction presented to participants prior to presenting items for assessment of the offences

For seriousness assessment: "Suçun birçok çeşidi vardır. Bazıları çok ağır sayılırken, bazıları çok ağır sayılmaz. Sonraki adımda değişik türde suçların bir listesini göreceksiniz. Biz, suçlardan hangisinin ne kadar ağır olduğu konusundaki fikirlerinizi merak ediyoruz.

Bir suçun hiç ağır olmadığını düşünüyorsanız, "0" işaretleyiniz.

Bir suçun en ağır olanlardan olduğunu düşünüyorsanız, o zaman "10" işaretleyiniz.

Bir suçun hiç ağır olmayan ile en ağır olanlar arasında bir yerde olduğunu düşünüyorsanız, o zaman 0 ve 10 arasında suçun ağırlığını en iyi temsil ettiğini düşündüğünüz sayıyı işaretleyiniz.

---

6   Original item 29 was not used in this study, as it refers to parking meters which at this time in Turkey do not exist, and it would have been a preposterous question.

7   Original items was "Shoplifting merchandise worth 600 from a store", however due to a typo in a survey a Turkish version item referred to "600TL", which in effect would have been equivalent to $60.

Unutmayın ki suçların ağırlığı bir şahsi görüş meselesidir ve biz sizin şahsi görüşünüzü merak ediyoruz."

*For wrongfulness assessment***:** "Şimdi, bir sonraki adımda her suç türü için, bir kişinin bu suçu işlemesinin sizin görüşünüzde ahlaken ne kadar yanlış olduğunu belirtiniz.

Eğer bir suçu işlemenin ahlaki açıdan hiç yanlış olmadığını düşünüyorsanız, "0" işaretleyiniz.

Eğer bir suçu işlemenin ahlaki açıdan çok yanlış olduğunu düşünüyorsanız, "10" işaretleyiniz.

Eğer değerlendirmenizde arada kaldıysanız, bir suçu işlemenin ahlaken ne kadar yanlış olduğu konusundaki düşüncenize denk düşen sayıyı işaretleyiniz."

For harmfulness assessment: "Son olarak listedeki suçların, mağdurlarına ne kadar zarar verdikleri, mağdurları ne kadar yaraladıkları hakkındaki düşüncelerinizi öğrenmek istiyoruz.

Eğer bir suçun mağdura pek de zarar vermediğini, onu yaralamadığını düşünüyorsanız, "0" işaretleyiniz.

Eğer bir suçun mağdura çok zarar veren veya onu çok yaralayan bir suç olduğunu düşünüyorsanız, "10" işaretleyiniz.

Eğer değerlendirmenizde arada kaldıysanız, bir suçun mağdur açısından ne kadar zarar veren bir suç olduğu konusundaki düşüncenize denk düşen sayıyı işaretleyiniz."

Annex III. Multiple choice question used to test whether the participant was aware of the criterion that needed to be used when assessing the items (same questions was used in all three stages of the survey)

"Bu aşamada suçları neye göre değerlendirmenizi istemiştik?

- Suçun ağırlığına göre

- Suçun ahlaki yanlışlığına göre

- Suçun niteliğine göre

- Suçun yarattığı zarara göre"

Annex IV. Instruction provided to those participate who provided the wrong answer the question in Annex III

For *seriousness* assessment stage of the survey: "Hayır. Suçlardan hangisinin ne kadar ağır olduğu konusundaki fikirlerinizi merak ediyoruz. Bir suçun hiç ağır olmadığını düşünüyorsanız, "0" işaretleyiniz. Bir suçun en ağır olanlardan olduğunu düşünüyorsanız, o zaman "10" işaretleyiniz. Bir suçun hiç ağır olmayan ile en ağır olanlar arasında bir yerde olduğunu düşünüyorsanız, o zaman 0 ve 10 arasında suçun ağırlığını en iyi temsil ettiğini düşündüğünüz sayıyı işaretleyiniz. Unutmayın ki suçların ağırlığı bir şahsi görüş meselesidir ve biz sizin şahsi görüşünüzü merak ediyoruz."

For *wrongfulness* assessment stage of the survey: "Hayır. Bu aşamada, her suç türü için, bir kişinin bu suçu işlemesinin sizin görüşünüzde ahlaken ne kadar yanlış olduğunu merak ediyoruz. Eğer bir suçu işlemenin ahlaki açıdan hiç yanlış olmadığını düşünüyorsanız, "0" işaretleyiniz. Eğer bir suçu işlemenin ahlaki açıdan çok yanlış olduğunu düşünüyorsanız, "10" işaretleyiniz. Eğer değerlendirmenizde arada kaldıysanız, bir suçu işlemenin ahlaken ne kadar yanlış olduğu konusundaki düşüncenize denk düşen sayıyı işaretleyiniz."

For *harmfulness* assessment stage of the survey: "Hayır. Son olarak listedeki suçların, mağdurlarına ne kadar zarar verdikleri, mağdurları ne kadar yaraladıkları hakkındaki düşüncelerinizi öğrenmek istiyoruz. Eğer bir suçun mağdura pek de zarar vermediğini, onu yaralamadığını düşünüyorsanız, "0" işaretleyiniz. Eğer bir suçun mağdura çok zarar veren veya onu çok yaralayan bir suç olduğunu düşünüyorsanız, "10" işaretleyiniz. Eğer değerlendirmenizde arada kaldıysanız, bir suçun mağdur açısından ne kadar zarar veren bir suç olduğu konusundaki düşüncenize denk düşen sayıyı işaretleyiniz."