

Hofmann, Elisa; Bolesta, Deliah; Adra, Aya

Working Paper

Immorality judgments and framing effects in voluntary payment settings

Jena Economic Research Papers, No. 2023-010

Provided in Cooperation with:

Friedrich Schiller University Jena, Faculty of Economics and Business Administration

Suggested Citation: Hofmann, Elisa; Bolesta, Deliah; Adra, Aya (2023) : Immorality judgments and framing effects in voluntary payment settings, Jena Economic Research Papers, No. 2023-010, Friedrich Schiller University Jena, Faculty of Economics and Business Administration, Jena

This Version is available at:

<https://hdl.handle.net/10419/283181>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Immorality Judgments and Framing Effects in Voluntary Payment Settings

Elisa Hofmann | Deliah Bolesta | Aya Adra

JENA ECONOMICS RESEARCH PAPERS · # 2023-010

Immorality Judgments and Framing Effects in Voluntary Payment Settings

Elisa Hofmann* Deliah Bolesta[†] Aya Adra[‡]

March 31, 2023

Abstract

The Theory of Dyadic Morality (TDM; Schein and Gray (2018)) posits that immorality judgments emerge from norm violations, harm perceptions, and negative affect. We test this core prediction in an applied setting: voluntary payment settings, such as the Pay-What-You-Want mechanism. In our study, we assess own payment intentions and how voluntary payments of an ostensible individual for an online-news website are judged by participants regarding their perceptions of immorality, harm, anger, and social norms. As political orientation is a key variable in theorizing and exploring immorality judgments in psychological research, we take its potential impact into account in our study. Because voluntary payments have been shown to be sensitive to framing, we vary the pricing mechanism's name in a between-subjects one-factorial design with four factor levels (*Pay-What-You-Want*, *You-Can*, *It-Is-Worth-To-You*, *You-Believe-Is-Fair*). The results of our online experiment with 602 Americans indicate that voluntary payment settings are indeed perceived as moral domains. We find that perceptions of norm violation, harm, and negative affect predict immorality judgments, lending empirical support to the Theory of Dyadic Morality. We also show that these components, the immorality judgments, and the own payment intentions are sensitive to framing effects. Finally, we find substantial differences between liberals and conservatives, suggesting an ideological influence on immorality judgments.

JEL classifications: C99; D01; D91; L11

APA classifications: 3000; 3040; 3920

Keywords: Theory of Dyadic Morality; immorality judgments; experiment; voluntary payments; Pay-What-You-Want; framing; social norms

We would like to thank the IMPRS Uncertainty fellows and the audience at IMEBESS2019 for their feedback. We are thankful to Uwe Cantner, Christoph Engel, Thomas Kessler, and Oliver Kirchkamp for their helpful comments. Financial support of the International Max Planck Research School "Adapting Behavior in a Fundamentally Uncertain World" is gratefully acknowledged. Ulrike Stoll provided excellent research assistance.

*Corresponding author, elisa.hofmann@uni-jena.de, Friedrich Schiller University Jena, Department of Intercultural Studies and Business Communications, Intercultural Human Resource and Organizational Development, Ernst-Abbe-Platz 8, D-07743 Jena, Germany

[†]deliah.bolesta@zfkfs.de, Center for Criminological Research Saxony, Karl-Liebknecht-Str. 29, D-09111 Chemnitz, Germany

[‡]aya.adra@esade.edu, Ramón Llull University, Esade Business School, Av. Pedralbes, 60-62, E-08034 Barcelona, Spain

“Morality, like art, means
drawing a line somewhere.”

Oscar Wilde

1 Introduction

The Theory of Dyadic Morality (TDM) (Schein and Gray, 2018) suggests that immorality judgments stem from a combination of negative affect, harm perception, and social norm violation. While the quantity of theoretical literature in the field of moral psychology regarding immorality judgments is increasing, empirical designs testing the Theory of Dyadic Morality remain scarce. It thus remains an open empirical question whether perceptions of social norms, anger, and harm perceptions predict immorality judgments, whether immorality judgments are independent of an individual’s political orientation, and whether such judgments are sensitive to framing effects in an applied setting: voluntary payment settings, such as Pay-What-You-Want pricing mechanisms.¹ Hence, we aim to examine whether such voluntary payment markets are morally laden. By using the Theory of Dyadic Morality as our theoretical framework, we explore whether immorality perceptions can be predicted by the three core elements of the Theory of Dyadic Morality, whether the framing of the voluntary payment setting has an impact on the intended payment behavior, the immorality judgments as well as on its three predictive elements, and what role an individual’s political orientation plays for immorality judgments.

Using an online experiment, we investigated how different levels of voluntary payments for an online-news website are evaluated. Essentially, our participants had to rate their perceptions of anger, harm, and social norms and indicate their judgment of immorality toward a series of choices made by an ostensible Individual A. In a between-subjects design, we exogenously varied the name of the voluntary payment setting to operationalize framing effects, resulting in a one-factorial design with four factor levels. The four framing treatments of the voluntary payment settings were based on previous research (Cui and Wiggins, 2017; Machado and Sinha, 2015; Saccardo et al., 2021; Sleesman and Conlon, 2017) and were called Pay-What-You-Want (PWYW), Pay-What-You-Can (PWYC), Pay-What-It-Is-Worth-To-You (PWIIWTY), and Pay-What-You-Believe-Is-Fair (PWYBIF). A total of 602 American test subjects participated in the experiment and were randomly assigned to one of the four framing treatments.

Our study has three main results. First, we find that perceptions of social norms and anger predict immorality judgments. Furthermore, we find that perceived harm consistently predicts immorality judgments beyond social norms and anger. We thus conclude that our study provides empirical evidence

¹Natter and Kaufmann (2015) refer to payments in such pricing mechanisms as “voluntary market payments”. We adapt this term and call markets which use such voluntary pricing mechanisms “voluntary payment markets”, as the market transactions themselves are characterized by voluntary payments. Throughout this paper we use the terms “voluntary payment markets”, “voluntary payment settings”, and “voluntary payment mechanisms” synonymously.

for the Theory of Dyadic Morality. Second, we find differences in the process of immorality judgments explained by political orientation. Conservatives' judgments of immorality are more strongly predicted by anger than harm, while the immorality judgments of liberals are most strongly predicted by harm. These results add to the research in moral psychology by providing further evidence that immorality judgments are sensitive to political orientation. Finally, our results demonstrate that the framing of the pricing mechanism does seem to have an effect on perceived harm, anger, social norms, immorality judgments, and payment intentions. Payment intentions are highest for the PWYBIF treatment. Further, while in all four treatments participants agree that low payments are more immoral and more socially inappropriate than high payments, in the PWIIWTY and PWYBIF treatments, immorality judgments are highest. This result indeed suggests that voluntary payment markets are perceived as a moral domain.

Our paper contributes to three streams of literature: Research on immorality judgments, research on voluntary payment markets including framing effects, and research on political orientation.

The last decades of research in social and moral psychology have witnessed the advancement of a multitude of theories attempting to account for how individuals make moral judgments. Although defining moral judgments has proven to be a challenging philosophical exercise, one widely used working definition is "*evaluations (good vs. bad) of the actions or character of a person* that are made with respect to a set of virtues held to be obligatory by a culture or subculture" (Haidt, 2001, 817). Different approaches have put forth diverging explanations for how people come to make such evaluations, including models of relationship regulation, which place social relations at the core of moral motives (Rai and Fiske, 2011), and modular models, which propose the existence of a number of independent moral domains in the mind (Haidt, 2012).

A recent development in the field has come with the advent of the Theory of Dyadic Morality (Schein and Gray, 2018), which regards moral judgment as revolving around a cognitive template of harm. This theoretical approach in explaining immorality judgments and its empirical test is the main focus of this study. Schein and Gray (2018) argue that moral judgments arise from a combination of three relevant and interrelated components, namely the perception of a norm violation, a negative affective reaction, and a perception of harm. In the literature it is assumed that norm violations are an essential precursor of immorality judgments. Defining norms as expectancies or beliefs about how others act and should act (Reno et al., 1993), the role of such expectancies and beliefs in predicting behavior and moral judgment (see, e.g., Kochanska et al., 2001) is a key aspect. Schein and Gray (2018) further acknowledge that different norms are more or less moralized and that negative affect could well be what moralizes norms. This fits well in the literature that has demonstrated the importance of negative affect in morality. Existing literature emphasizes the consistent finding that emotions, particularly negative emotions, are crucial for moral cognition (see, e.g., Valdesolo and DeSteno, 2006). Specifically, studies have found that

immorality judgments are often laden with negative affect (Haidt, 2001) and experimental setups have shown that increasing negative affect can lead to an increase in immorality judgments (see, e.g., Horberg et al., 2011; Inbar et al., 2012). Harm perceptions are advanced by Schein and Gray (2018) as the third required component. This assumption is to highlight the fact that ‘wrongness’ as it pertains to morality, arises from norm violations that cause negative affect, only insofar as these norm violations engender perceptions of harm (Schein and Gray, 2018).

Previous research, anchored in different domains and using different methodologies, has consistently highlighted a strong relationship between harm perceptions and immorality judgments (see, for a review, Schein and Gray, 2018). For example, the condemnation of acts that are as different as incestuous sex and genetically modified organisms (GMOs), has been demonstrated to be significantly predicted by harm perceptions (see, e.g., Gray and Schein, 2016; Royzman et al., 2015). Furthermore, Schein and Gray (2015) showed that when asked to freely recall acts that are morally wrong, participants overwhelmingly described situations that they perceived as harmful. Consistent with this finding, Schein and Gray (2015) also showed that participants judged scripts which involved harm significantly more harshly than those that did not. Perhaps most importantly, experimental manipulations of the presence of harm have suggested a causal link between harm perceptions and the severity of immorality judgments (Schein and Gray, 2015). The TDM links these findings with the assumption that harm might dominate everyday morality, the law, and psychological findings with what predicts people’s evaluations of immorality.

Put together, the synthesis of these research traditions suggests a triadic conception of moral judgments consisting of norms, negative affect, and harm perceptions, whereby perceived harm is assumed to be the fundamental predictor of immorality judgments (Schein and Gray, 2018). Schein and Gray (2018) contend that “morality emerges from the overlap of norms, affect, and harm perceptions, with perceived harm itself emerging from the overlap of agents, patients, and causation” (p. 56). Thus, all three core components – anger, social norm violation, and harm perceptions – seem to be essential in leading to an immorality judgment.

Importantly, within the framework of the Theory of Dyadic Morality, perceived harm is dyadic by nature, meaning that it involves the perception of an intentional moral agent that causes damage to a vulnerable moral patient (Schein and Gray, 2018). A moral agent is defined as someone having the “capacity to do right or wrong” (Gray and Wegner, 2009, p. 505), whereas a moral patient is defined as someone having the “capacity to be a target of right or wrong” (Gray and Wegner, 2009, p. 505). As Gray and Wegner (2009) clarify, the definition of who is a moral agent and who is a moral patient is intuitive, subjective and “perceived as matters of degree” (p. 506), whereas the process of categorizing an individual as moral agent or moral patient is called “moral typecasting” (Schein and Gray, 2018). Schein and Gray (2016) call this underlying cognitive principle “harm-based dyadic template” (p. 62). Accordingly, they assume that the perception of harm is also subjective, intuitive and thus “in the eye

of the beholder” (Schein and Gray, 2016, p. 39). Schein and Gray (2018) further propose that the perception of harm and immorality is not dichotomous but happens on a continuum.

To the best of our knowledge, the Theory of Dyadic Morality, while theoretically compelling and widely discussed (see, e.g., Gray and Graham, 2019), has never been tested directly and systematically in an empirical manner. The current paper aims at filling this gap and contributing to research on moral psychology by mapping and testing the theoretical predictions of Schein and Gray’s Theory of Dyadic Morality in an applied setting, namely, voluntary payment settings. Our empirical approach allows us to systematically investigate the roles of norms, anger (a negative affect), and harm perceptions in predicting immorality judgments in this realistic context of voluntary payment markets.

Previous research has expressed the fear that markets erode moral values (see, e.g., Bartling et al., 2023; Storr and Choi, 2019; Falk and Szech, 2013). Voluntary payment markets, such as Pay-What-You-Want settings, can be interpreted as special types of markets. In these voluntary payment settings, the whole price determination is transferred to the buyers (Kim et al., 2009). More and more organizations apply voluntary payment mechanisms, for instance the German newspaper ‘Die Tageszeitung’ (TAZ), the newspaper ‘The Guardian’, the restaurant ‘Wiener Deewan’ (Vienna, Austria), the ‘American Museum of Natural History’ (New York, USA), the online-shop ‘Yummy Organics’, the hotel ‘ibis Mumbai Vikhroli’ (Mumbai, India) as well as the ‘Theater Tübingen’ (Tübingen, Germany). Voluntary payment markets have been explored in the fields of behavioral economics and marketing in recent years (see, for recent reviews, Gerpott, 2017; Greiff and Egbert, 2018; Vizuite-Luciano et al., 2022). Various studies demonstrate that people pay positive amounts in Pay-What-You-Want contexts (Gneezy et al., 2010; Kim et al., 2009; Krämer et al., 2017; Regner and Barria, 2009; Schmidt et al., 2015). Previous research focused on two aspects, namely the underlying motivations of individuals for paying a positive amount voluntarily, also called personal-related factors, and external success factors for a voluntary payment scheme, also called situational factors (Kunter, 2015). Personal-related factors are such as social preferences, internal reference prices, loyalty, and price consciousness (Greiff and Egbert, 2016; Kim et al., 2009; Kunter, 2015; Natter and Kaufmann, 2015; Regner and Barria, 2009; Schmidt et al., 2015). Regarding the situational factors, variables such as external reference prices, knowledge about the product, and social norms have been established as relevant drivers of the payment decision (Armstrong Soule and Madrigal, 2015; Greiff et al., 2013; Jang and Chu, 2012; Mak et al., 2015; Narwal and Nayak, 2019; Riener and Traxler, 2012).

The change in price determination power in voluntary payment markets on the one hand increases perceptions of price fairness (Haws and Bearden, 2006) and also allows buyers to pay more than fixed prices. On the other hand, voluntary payment markets come along with a risk for the sellers: buyers alone decide how much to pay for a market transaction. This implies that buyers who pay nothing or

only low amounts cannot be punished for paying less. Hence, voluntary payments allow for both egoistic and social preferences. Voluntary payment markets can, moreover, naturally be seen as scenarios in which the morality of a certain behavior (i.e. payment) is relevant. In the light of the TDM, buyers in voluntary payment settings can be seen as the moral agents. They are able to do right or wrong depending on the price they pay. The seller, on the other hand, can be seen as the moral patient. They are not able to reject payments and thus are able to suffer from payments of the buyer that are too low. This asymmetric relationship further results in high uncertainty for the buyer regarding the price. In voluntary payment contexts without a suggested price it might not be clear which payment behavior is moral and which is immoral. Also, social norm perceptions remain uncertain for the buyer in voluntary payment settings (Feldhaus et al., 2018), when the expectations of the seller and the behavior of other buyers is taken into account. Hence, it is still unknown whether voluntary payment markets are perceived as a moral domain by the buyers or whether social norms shape such voluntary payment markets.

As voluntary payment settings are real existing settings in which it is unclear whether harm, anger, and a social norm violation are perceived, they are externally more valid compared to previously used examples of testing theories of immorality, such as cleaning the toilet with a national flag (Haidt et al., 2000). Using such real world contexts like voluntary payment settings avoids building a “platypus moral psychology” based on atypical behavior (Schein and Gray, 2018). Furthermore, they represent settings of ‘everyday morality’ which allow moral wiggle room (Dana et al., 2007) for consumers if they behave self-interestedly and pay low amounts. However, it is unclear yet if this is the case or whether consumers nevertheless unanimously have similar perceptions of immorality, harm, anger, and social norm violations regarding different amounts of prices paid in voluntary payment settings. Being such a subliminal setting of daily (im)morality, voluntary payment markets are an interesting and critical environment to test the Theory of Dyadic Morality.

First, we contribute to the existing research on voluntary payment mechanisms by examining how such markets are perceived regarding their immorality and social inappropriateness. Furthermore, we contribute to the research in moral psychology by empirically testing the Theory of Dyadic Morality in an applied setting. In summary, we bridge the two disciplines of moral psychology and behavioral economics by applying a theoretical framework drawn from the first onto a commonly investigated mechanism from the second.

Second, we contribute to the literature on framing effects in voluntary payment markets. The uncertainty in voluntary payment settings regarding pricing leaves room for another concept potentially influential on the payment decision: framing effects. In the literature, ‘framing effects’ usually refers to the presentation of the same situation (equivalent regarding outcome) with different labels (Frisch, 1993; Gerlach and Jaeger, 2016; Tversky and Kahneman, 1981, 1986). One of the first examples of framing ef-

fects is presented in the concept of prospect theory of Kahneman and Tversky (1979), who detected that individuals react differently to the same gamble when labeled (framed) as losses or as gains. Framing effects are also known as the contextual cues of a situation (Brañas-Garza, 2007; Frisch, 1993), where the framing is assumed to “focus people’s attention on different attributes” (Frisch, 1993, p. 423). Extensive research has shown that framing effects exist in a broad variety of settings (see, e.g., Andreoni, 1995; Brewer and Kramer, 1986; Engel and Rand, 2014; Liberman et al., 2004; Korenok et al., 2014; Krupka and Weber, 2013; Sonnemans et al., 1998). Furthermore, as proposed by Frisch (1993) as well as Hoch and Ha (1986), framing effects might affect behavior, especially in those situations in which ambiguity about the appropriate behavior is high.

Several publications have appeared in recent years documenting the existence of framing effects in the context of voluntary payments, resulting in different payment behaviors. These studies indicate that voluntary payments are sensitive to contextual cues and that it thus matters how the pricing mechanism is communicated to the buyer. Framing effects in those studies have been operationalized as the name of the pricing mechanism. While the pricing mechanism itself stays the same (the buyer has complete control over the price paid), the buyers might perceive the mechanism differently according to the respective label. Recent studies have indicated that individuals pay more if the pricing mechanism is called Pay-What-You-Can (PWYC) compared to Pay-What-You-Want (PWYW) (Cui and Wiggins, 2017; Saccardo et al., 2021). In a similar way, Machado and Sinha (2015) as well as Cui and Wiggins (2017) compared the label Pay-What-It-Is-Worth-To-You (PWIIWTY) with the label Pay-What-You-Want (PWYW). They found that payments are higher if the pricing mechanism is called PWIIWTY than if it is called PWYW. Finally, a study by Sleesman and Conlon (2017) concluded that payments are higher when the pricing mechanism is called Pay-What-You-Believe-Is-Fair (PWYBIF) compared to PWYW. We take this into account as we study different framings of voluntary payment markets and their impact on payment behavior intentions as well as on immorality judgments, perceived harm, anger, and social norms. In this study, we focus on the following four framings of voluntary payment contexts, which have been identified as most common and investigated ones in prior research (Gerpott, 2017): Pay-What-You-Want (PWYW), Pay-What-You-Can (PWYC), Pay-What-It-Is-Worth-To-You (PWIIWTY), and Pay-What-You-Believe-Is-Fair (PWYBIF).

Taking the literature on framing effects and voluntary payments into account, we expect the different framings to create contexts in which different rules of immorality, anger, perceived harm, and social norm perceptions are activated, which in turn drive payment behavior intentions. We argue that the framing of the pricing mechanism is important as buyers might perceive the name as advice from the seller regarding payment behavior (Cui and Wiggins, 2017). Pay-What-You-Want can serve as baseline in which the adequate behavior would be to pay whatever an individual wants. Pay-What-You-Can might activate a behavior in which payments are dependent on the buyer’s available budget. Pay-What-It-Is-

Worth-To-You, on the other hand, might activate the perception that satisfaction with the consumption should influence the payment. Pay-What-You-Think-Is-Fair might activate a fairness norm, in which it is reasonable to pay a high price in order to behave fairly towards the seller. In this case it would be unfair to pay nothing as the seller incurs production costs. We expect that PWYBIF to be the strongest framing treatment of the four as it can be perceived as the most morally laden label.

To the authors' best knowledge, there is currently no empirical work comparing the effects of all four of the above-mentioned framings, neither on differences in payments nor on the differences in social normativity, harm perceptions, anger, and perceptions of immorality. Despite the importance of framing effects, there remains a paucity of evidence about whether the four different labels of the voluntary pricing mechanism result in different social norms, anger, harm, and immorality perceptions or in different payment behavior intentions. In particular, the existence of social norms in voluntary payment contexts is unexplored. Although various studies emphasize the relevance of social norms for the payment decision (Armstrong Soule and Madrigal, 2015; Jang and Chu, 2012; Narwal and Nayak, 2019), it remains an open empirical question which social norms are at play in those settings. Our study contributes to the existing research by testing whether the four framings used bring about different payment behavior intentions and different judgments of immorality, including different perceptions of social norms, harm, and anger.

The goal of this paper is twofold regarding framing effects: First, we want to investigate whether the name of the pricing mechanism (as a frame) influences payment behavior intentions as well as immorality, anger, perceived harm, and social norm perceptions. Differences are suggested by the existing literature but the four framings have not been explored in combination yet. Second, we use the four framings as robustness checks for testing the predictions of the Theory of Dyadic Morality since they are expected to hold in all four framing treatments.

Third, as political orientation is a key variable in explaining immorality judgments in moral psychology (Graham et al., 2009), we contribute to the literature on political orientation by taking its impact on immorality judgments into account. Identifying what drives an individual's moral judgment has a long-standing tradition within the social sciences. Graham et al. (2009) argue that differences in moral judgments are informed by political ideology. In their understanding, moral foundations are innate (Marcus, 2004) and thus, are "organized in advance of experience" (Graham et al., 2009, p. 1031). Relying on the assumption that an either conservative or liberal ideology relates to a specific, often complementary, structured personality (Jost et al., 2003), they assume conservatives and liberals to rely on a different set of moral foundations when making moral judgments.² Conservatives were repeatedly found to endorse each moral foundation equally whereas liberals more strongly endorsed those related to harm and fairness.

²The five moral foundations according to Graham et al. (2009) are harm/care, fairness/reciprocity, ingroup/loyalty, authority/respect, and purity/sanctity.

In a more recent approach, Schein and Gray (2015) contest the assumption of there being five moral foundations and suggest the concept of dyadic morality with the central foundation being harm. In their studies they found that conservatives and liberals both perceived harmful acts as more immoral than unfair, disloyal, disobedient or impure acts. Except for the impurity domain (for which the effect could not be replicated in subsequent studies) there was no interaction between political orientation and the moral foundations. They argue that differences in moral judgment between conservatives and liberals may occur due to the fact that they “legitimately differ on specific issues” (p. 1160) but not necessarily because of fundamental, essentialist differences in moral reasoning. Thus, if moral judgment is inferred from a set of stimuli that is *known* to show political differences (Haidt, 2012), drawing inferences about fundamental psychological processes may be erroneous. In general, the Theory of Dyadic Morality assumes that liberals and conservatives might differ regarding what they perceive as harmful and immoral but that “liberals and conservatives share fundamentally the same moral mind” (Schein and Gray, 2018, p. 35). Likewise, Gray et al. (2014) have “found substantial similarity between liberals and conservatives” (p. 9). Given voluntary payments focusing on prices paid occur within an economic setting, it is all the more important to account for potential influences of political orientation on immorality judgments. This approach takes into account existing literature that emphasizes the potential domain specificity of political cognition (see, e.g., Brandt et al., 2014; Conway et al., 2016; Fiagbenu et al., 2019; Morgan et al., 2010; Proch et al., 2019). As this study aims to test the predictions of the TDM, we focus on this theory in the sections following and assume that the core predictions of the theory are valid irrespective of political orientation.³ However, we take the potential impact of political orientation into account to control for the possibility that the process of (im)morality judgments might be sensitive to political orientation.

The remainder of the paper is structured as follows: Section 2 describes the experimental design while Section 3 is devoted to the hypotheses. Section 4 focuses on the experimental procedures and participants of the study. The results are presented in Section 5, followed by a discussion in Section 6, with Section 7 concluding the paper.

2 Experimental design

By means of an online experiment we explored how different levels of voluntary payments for an online-news website were evaluated. In particular, subjects had to rate their perceptions of anger, harm, and social norm as well as their immorality judgment towards a series of behaviors made by an ostensible Individual A. We used this third-party spectator design in order to reduce the bias towards one’s own

³We empirically test the predictions of the TDM based on its own claims. However, we are aware that there are important conceptual limitations that might be worth discussing, e.g. whether other plausible factors exist that could predict immorality judgments beyond the three components of TDM. While such a conceptual critique is not the focus of this study, we discuss potential areas for further research on this topic in Section 6.

(immoral) behavior. The three core predictive components were operationalized as measures of anger, global harm perception, and an adapted version of the Krupka and Weber (2013) social norm elicitation, respectively. The Krupka-Weber elicitation method of social norms uses coordination on the anticipated social norm to measure them as “collective perceptions” (Krupka and Weber, 2013, p. 499). Thus, this method captures injunctive social norms, which refers to social norms conceptualized as the behavior people believe others should show. Using a point measure on a 4-point-scale ranging from “very socially inappropriate” (1) to “very socially appropriate” (4), we assume that this approach elicits differences in social norm perceptions.⁴

The framing of the voluntary payment markets, henceforth referred to as ‘treatment’, was varied between subjects, resulting in a one-factorial between-subjects design with four factor levels. Table 1 contains a summary of the treatments and the observations per treatment. The framings of the four different voluntary payment settings were called Pay-What-You-Want (PWYW), Pay-What-You-Can (PWYC), Pay-What-It-Is-Worth-To-You (PWIIWTY), and Pay-What-You-Believe-Is-Fair (PWYBIF).⁵

Treatment	Observations
Pay-What-You-Want (PWYW)	155
Pay-What-You-Can (PWYC)	146
Pay-What-It-Is-Worth-To-You (PWIIWTY)	151
Pay-What-You-Believe-Is-Fair (PWYBIF)	150
<i>Total</i>	602

Table 1: Treatments and observations

As Schein and Gray (2018) claim that both, harm and immorality perceptions, exist along a continuum, we designed scenarios in which harm and immorality perceptions were able to exist along such a continuum. We chose six different payments in \$ (\$0, \$1, \$2, \$3, \$4, \$5) for an online-newspaper with each price paid possibly resulting in a more or less salient amount of harm perceptions and immorality.⁶ This range from \$0 to \$5 was chosen on the basis of the weekly price range of digital newspapers in the USA in 2017 (Watson, 2018). All participants had to evaluate all six possible prices paid, representing six

⁴However, we are aware that the Krupka-Weber social norm elicitation method captures injunctive social norms only without claiming to be able to differentiate between other principles of social or distributive justice, such as equity, equality or need (see, e.g., Deutsch, 1975; Jost and Kay, 2010; Rawls, 1971).

⁵The framing selection was the result of a pretest ($N = 231$) run in October 2018. Participants took an average of $M = 6.6$ minutes ($SD = 4.15$) to complete the pretest survey and were paid \$0.50 for their participation. The four above mentioned and three further framings (Pay-What-You-Like, Pay-What-You-Will, Pay-What-You-Wish) were tested for their respective associations. The framings differed neither significantly nor systematically in their respective associations in the pretest. Due to semantic equivalence of the verbs ‘want’, ‘like’, ‘will’, and ‘wish’, we did not expect these four framings to differ with regard to the social norm perceptions elicited. Thus, we did not expect that including the four similar framings would contribute to our theory testing substantially and decided to include Pay-What-You-Want (PWYW) as a baseline treatment only. The four selected framings PWYW, PWYC, PWIIWTY, and PWYBIF were identified as most common and investigated ones in prior research (Gerpott, 2017). We expected them to differ regarding the social norm perceptions elicited and possibly also regarding perceived harm, anger, and immorality perceptions. Thus, we decided to use Pay-What-You-Want, Pay-What-You-Can, Pay-What-It-Is-Worth-To-You, and Pay-What-You-Believe-Is-Fair as the four framing treatments in this study while being aware that this is a limited selection of potential voluntary payment framings.

⁶We use the terms ‘payment(s) in \$’, ‘paid amount(s)’, and ‘price(s) paid’ interchangeably in this paper.

within-subjects measurements per participant in our design. The six prices were presented in ascending order from \$0 to \$5. In our study, paying \$0 was considered a situation in which it may be relatively easy to detect immoral behavior, whereas this may not be the case in the paying \$5 situation. We thus aimed to manipulate harm perceptions and immorality ratings with our six payments in \$ in order to explore whether immorality judgments, perceived harm, anger, and social norm perceptions differ for low levels of payment or high levels of payment. Consuming a product without paying for it financially harms the seller and might in turn lead to a shutdown of the business. Hence, paying \$0 can be seen as a situation which is perceived as both, harmful and immoral. Paying \$5 can be seen as a situation which is perceived as neither harmful, nor immoral. We chose this study design due to the logical connection between the order from low to high payments.⁷

3 Hypotheses

The main goals of this study are to empirically test the Theory of Dyadic Morality in the realistic setting of voluntary payment markets and to examine whether voluntary payment markets are perceived as a moral domain. We further test whether the process of immorality judgments is sensitive to political orientation. Finally, we investigate whether perceptions of immorality, harm, anger, and social norms as well as own payment intentions are sensitive to the respective framing of the voluntary payment setting. Accordingly, our hypotheses are grouped along these three categories: Testing the Theory of Dyadic Morality, influence of political orientation, and framing effects.

The key proposition of the Theory of Dyadic Morality is that perceived immorality is predicted by perceptions of harm, anger, and social norms. Perceived harm – in this model – is assumed to be the most fundamental and important predictor (Gray and Schein, 2016; Schein and Gray, 2018). Referring to Schein and Gray (2018), the Theory of Dyadic Morality would be falsified, if “when controlling for norms and affect – intuitive perceptions of harm are not causally linked to intuitive perception of immorality” (p. 58). Given the above argumentation, we hypothesize that:

Hypothesis 1 *Perceptions of social norms will positively predict immorality judgments.*

Hypothesis 2 *Anger will positively predict immorality judgments.*

Hypothesis 3 *Harm perceptions will positively predict immorality judgments, beyond perceptions of social norms and anger.*

Political orientation is a central variable in explaining immorality judgments in moral psychology. Schein and Gray (2018) propose that the underlying mechanism of harm and immorality perception is

⁷The design was tested in a pilot study ($N = 195$, after removing outliers and cases where attention checks were failed) in November 2018. Participants took $M = 9.98$ ($SD = 6.65$) minutes to complete the survey and were paid \$1.50 for their participation. We ran two one-factorial ANOVAs with immorality or harm perception as dependent variable and the payment in \$ as predicting factor. Planned contrasts revealed that paying \$0 was perceived as being significantly more immoral ($F(1, 2134) = 46.92, p < .001$) and more harmful ($F(1, 2134) = 118.20, p < .001$) than paying \$5.

similar irrespective of political orientation. Accordingly, the expectation is that immorality and harm perceptions follow the same process for liberals and conservatives. Hence, we predict the following:

Hypothesis 4 *The process of immorality judgments is similar for both, liberals and conservatives. Thus, perceptions of harm, social norm, and anger predict immorality judgments irrespective of political orientation.*

Voluntary pricing mechanisms, framed in various ways, may convey different information on social norms and result in differences regarding negative affect, perceived immorality, and harm perceptions. Wordings like ‘worth’ or ‘fair’ in the pricing mechanism name can be expected to trigger different expectations regarding behavior than ‘want’ or ‘can’. More generally speaking, it might be the case that ‘worth’ or ‘fair’ framings represent contexts which are more morally laden than ‘want’ or ‘can’ framings. It is currently unknown whether perceptions of immorality, harm, anger, and social norms are framedependent. We assume that, although the underlying pricing mechanism is identical in all treatments, the framing of the mechanism leads to different perceptions of the market environment. Previous research further suggests an impact of framing on payment behavior, namely that a framing as PWYC, PWIIWTY or PWYBIF increases payments compared to PWYW. We further expect that PWYBIF leads to higher payments than PWIIWTY and PWYC. Given the above argumentation regarding framing, we hypothesize that:

Hypothesis 5 *The four voluntary payment frames activate different perceptions of social norms, harm, anger, and immorality.*

Hypothesis 6 *The four voluntary payment frames lead to differences in the Subject’s payment behavior intention: Pay-What-You-Believe-Is-Fair leads to the highest own payment behavior intention, followed by Pay-What-It-Is-Worth-To-You, Pay-What-You-Can and Pay-What-You Want.*

4 Procedures and participants

After welcoming the participants to the experiment, in **Step 1** the subjects were introduced to the scenario description.⁸ Participants were randomly assigned to one of the four framing treatments. Participants were under the impression that there exists an online-newspaper called ‘Today’s Globe’ which offers a specific pricing mechanism for the consumption of their online-news.⁹ In the PWYW treatment, the pricing mechanism was called Pay-What-You-Want; in the PWYC treatment it was called Pay-What-You-Can; in the PWIIWTY treatment it was called Pay-What-It-Is-Worth-To-You; and in the PWYBIF

⁸The instructions are provided in the Appendix. All study materials, data, and analysis scripts are available on OSF.

⁹The ‘Today’s Globe’ is a fictitious online-news site which does not exist in reality. Although we are aware about the discussion in experimental economics on the use of deception (Charness et al., 2021), we decided to use such a hypothetical news-outlet to avoid potential biases and attitudes towards online-newspapers, especially with respect to political orientation. We interpret the reliability of the payment intentions in the discussion. However, we assume that this study design does not systematically affect differences in payment intentions or perceptions of immorality, social appropriateness, anger, and harm between the treatments due to randomization of the participants.

treatment the pricing mechanism was called Pay-What-You-Believe-Is-Fair. The subjects learned that the readers can choose on their own how much to pay for their weekly consumption of online-news at ‘Today’s Globe’. A screenshot of the fictitious website was shown to the participants emphasizing the newspaper’s independence and high-quality journalistic standards. After answering a control question about the name of the ‘Today’s Globe’ pricing mechanism the subjects entered Step 2.

In **Step 2** participants were asked to evaluate six different behavior options (within-subject measurement) of an ostensible Individual A who reads the online-news website of ‘Today’s Globe’. It was stated that Individual A was satisfied with the news they get at the online-news website of ‘Today’s Globe’. Participants then rated a series of six payment behaviors (ranging from paying \$0 to paying \$5 as weekly payment) ostensibly done by Individual A.

Subjects were asked to judge the payment behavior regarding perceived harm, anger, social norms, and perceived immorality (in random order). The measure for the social norm rating was adapted from Krupka and Weber (2013). Subjects rated the social appropriateness (ranging from “very socially inappropriate” (1) to “very socially appropriate” (7)) of six different prices paid and were also asked to provide what they think would be a ‘common’ rating rather than their ‘personal’ rating to capture the common sense characteristics of social norms. Anger was also measured for all six prices paid. Subjects had to state how angry each of the actions made them feel (ranging from “not angry at all” (1) to “very angry” (7)). For perceived immorality, subjects indicated how immoral they perceived the six prices paid options to be (ranging from “not immoral at all” (1) to “very immoral” (7)). Finally, the measure of harm was elicited as global harm and as dyadic harm. Global harm was measured by asking participants how harmful they perceived the six different prices paid of Individual A to be (ranging from “not harmful at all” (1) to “very harmful” (7)). The dyadic harm component was operationalized as an evaluation of the extent to which subjects agree or disagree with four statements, capturing the possibility of the buyer being the intentional agent and the seller being the victim of harmful actions and vice versa.

Step 3 contained various control measures, including own payment behavior intention¹⁰, assessment of a fair payment, and online usage.

The experiment ended in **Step 4** with a question about participants’ political orientation (self-placement on a scale ranging from “Strongly Liberal” (1) to “Strongly Conservative” (7)), and further demographics, including age, gender, education, employment status, income, marital status, and religiosity.

602 subjects participated in the experiment. They were recruited via the online platform Amazon Mechanical Turk (MTurk) and had to pass several qualifying criteria.¹¹ Participants were incentivized

¹⁰Subjects were asked to indicate the amount of USD they would be willing to pay weekly for the consumption of online-news, assuming they were satisfied with the quality of the news provided. This measure was counterbalanced by asking participants before or after measuring the other dependent variables, to control for order effects.

¹¹The qualifying criteria on Amazon MTurk were: Being U.S. citizen, 95% approval rate of previous Human Intelligence

with \$1.25 and took $M = 12.99$ minutes ($SD = 7.6$) to complete the experiment. The percentage of female subjects was 53% and the mean age was 39.8 years ($SD = 12.80$). A further 43% of the subjects identified as religious and the mean political orientation of our sample was 3.59 ($SD = 1.76$). The mean own payment intention was \$3.17 ($SD = 3.07$) with an evaluated fair payment being at a similar level ($M = \$3.16$, $SD = 2.47$), as indicated by a two-sample t-test ($t(1143) = 0.077$, $p = .939$).

5 Results

This paper aims at testing the predictions suggested by the Theory of Dyadic Morality and whether voluntary payment markets are morally laden. It was crucial that our design enabled us to induce acts varied by their subjectively perceived immorality, harm, anger, and social norms. We assumed that paying \$0 represents a situation in which both immorality and harm would be rather easily recognized by participants.

As illustrated in Figure 1, this assumption is confirmed across treatments, with participants perceiving paying \$0 as the most immoral and most harmful payment (see Table A.1 in the Appendix for descriptive statistics on perceptions of immorality, global harm, anger, and social norm of all six prices paid across treatments).

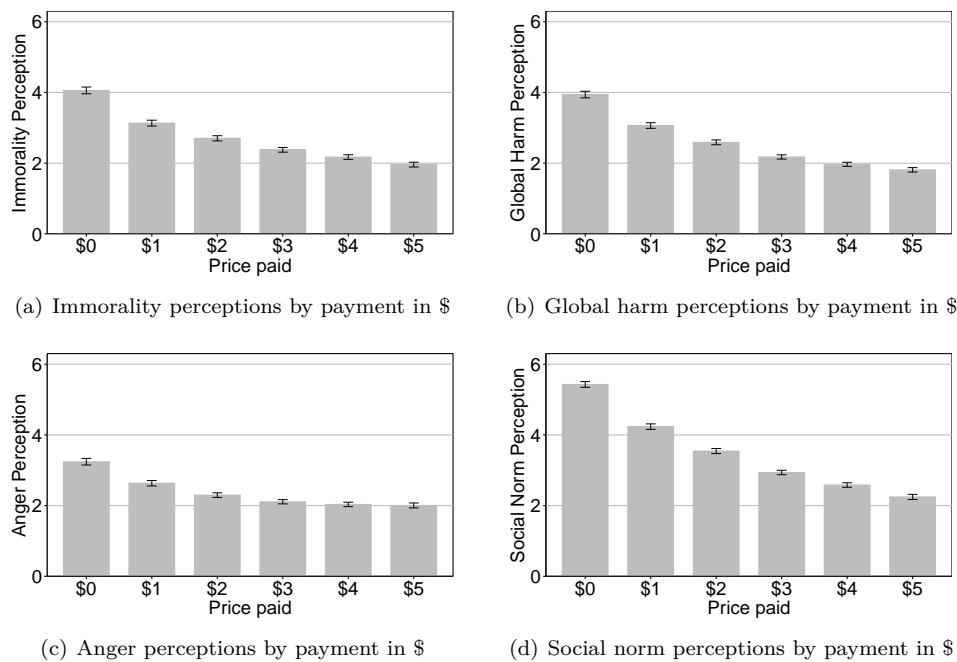


Figure 1: Perceptions of immorality, global harm, anger, and social norms by payment in \$

Immorality judgments and harm perceptions are highest for paying \$0; the same result applied to anger and social norm perceptions. Perceptions of immorality, global harm, anger, and social norms decrease with increasing payments and reach their respective minimum at \$5. Thus, the subjects agree

Tasks (HITs), and having more than 5000 HITs completed.

that paying \$0 and \$1 is very socially inappropriate behavior whereas paying \$4 and \$5 is very socially appropriate. Nonparametric Kruskal-Wallis tests show that immorality perceptions differ significantly across the six prices paid ($p < .001$). Also, differences in global harm, anger, and social norms between the six payments in \$ are statistically significant (Kruskal-Wallis tests, $p < .001$).

In addition, we ran regressions (mixed effects models) with immorality, global harm, anger, and social norms as dependent variables, respectively, and the six prices paid as predictor variable. Framing treatment was included as a random effect in each model. Regression results (see A.2 in Appendix) indicate that immorality judgments decreased significantly with increasing payments (coefficient = -0.39, $p < .001$, AIC = 14781.87, $R^2_{\text{LMM(m)}} = .11$, $R^2_{\text{LMM(c)}} = .12$) and that also global harm perceptions significantly decrease when payments increase (coefficient = -0.41, $p < .001$, AIC = 14387.00, $R^2_{\text{LMM(m)}} = .14$, $R^2_{\text{LMM(c)}} = .14$). We find the same results for anger (coefficient = -0.23, $p < .001$, AIC = 14389.93, $R^2_{\text{LMM(m)}} = .05$, $R^2_{\text{LMM(c)}} = .05$) and social norms (coefficient = -0.61, $p < .001$, AIC = 14443.55, $R^2_{\text{LMM(m)}} = .26$, $R^2_{\text{LMM(c)}} = .26$). These results allow us to assume we successfully designed an experimental environment to test the predictions of the Theory of Dyadic Morality.

One element in the Theory of Dyadic Morality is the concept of dyadic harm: the perception of a moral agent causing intentional harm to a vulnerable patient. In our voluntary payment setting we expect that the buyer is predisposed to be the moral agent (as they can cause financial harm to sellers if paying \$0) and the seller is predisposed to be the moral patient (as they can suffer from financial harm if the buyer pays low prices). Our data (see Table 2) lends support to this assumption. Buyers are perceived significantly more likely to be the moral agent ($M = 3.82$, $SD = 2.14$) than the moral patient ($M = 3.20$, $SD = 2.15$, $t(1202) = 5.05$, $p < .001$). In contrast, sellers are perceived significantly more likely to be the moral patient ($M = 4.11$, $SD = 2.10$) than the moral agent ($M = 2.96$, $SD = 2.08$, $t(1202) = 9.51$, $p < .001$).

	Agent Buyer	=	Patient Buyer	=	Agent Seller	=	Patient Seller	=	<i>n</i>
Whole dataset	3.82		3.20		2.96		4.11		602
Liberal (1-3)	3.88		3.15		2.95		4.24		284
Moderate (4)	3.71		3.16		2.95		3.91		141
Conservative (5-7)	3.82		3.30		2.99		4.06		177
<i>p value (Kruskal-Wallis test)</i>	.550		.672		.984		.320		

Table 2: Mean values of moral typecasting for the whole dataset and split by political orientation groups

Moral typecasting processes (categorizing who is a moral agent and who is a moral patient) were expected to be similar for both liberals and conservatives: they were investigated using a split sample where participants were categorized along their indicated political orientation as either ‘Liberal’ (anyone who stated a value of 1-3 on the self-placement scale of political orientation), ‘Moderate’ (anyone who

stated a value of 4), or ‘Conservative’ (anyone who stated a value of 5-7). We compared the answers to all four dyadic harm perception questions between these political categories. Our data, summarized in Table 2, suggests that there are no significant differences between political groups. Instead, we can see that neither categorization as moral agent nor as moral patient differs significantly between political groups (p values in Table 2 as obtained in nonparametric Kruskal-Wallis tests between the three political orientation categories, respectively). Moral typecasting processes thus seem to work similarly for conservatives, moderates, and liberals.

5.1 Testing the Theory of Dyadic Morality

In the first step, we looked at correlations between immorality judgments and perceptions of harm, anger, and social norms. Immorality judgments are slightly positively correlated with social norms (Pearson’s $r = 0.35$), highly positively correlated with global harm perceptions ($r = 0.61$), and highly positively correlated with anger ($r = 0.62$). All correlations are significant at the 1 % level.

In the second step, we ran regressions to test how perceptions of social norms, anger, and harm predict immorality judgments (see Table 3). Taking into account the nested structure of the data, we ran regressions (mixed effects models) with the three predictors (global harm perceptions, anger, social norm) as fixed effects, and the six payments in \$ nested within the four framing treatments as random effects. Immorality judgment is the dependent variable in all models. We first included global harm perceptions (Model 1), then added anger (Model 2) and finally social norm (Model 3) as independent variable, respectively. All three predictors are found to positively predict immorality judgments. When including all three predictors (Model 3), anger (coefficient = 0.37, $p < .001$) is found to be the strongest predictor of immorality judgments rather than global harm perceptions (coefficient = 0.36, $p < .001$).

When comparing the three regression models, Model 2 (AIC = 12802.49, $R^2_{\text{LMM(m)}} = .45$, $R^2_{\text{LMM(c)}} = .47$) fits the data significantly better than Model 1 (AIC = 13324.98, $R^2_{\text{LMM(m)}} = .37$, $R^2_{\text{LMM(c)}} = .39$, $\chi^2 = 524.48$, $p < .001$). Model 3 (AIC = 12739.39, $R^2_{\text{LMM(m)}} = .48$, $R^2_{\text{LMM(c)}} = .49$) fits the data significantly better than Model 2 ($\chi^2 = 65.10$, $p < .001$, see Table 4 for full ANOVA).

Result 1 *All three core elements of the Theory of Dyadic Morality (anger, social norms, and harm perceptions) positively and significantly predict immorality judgments. Harm predicts immorality perceptions beyond perceptions of anger and social norms.*

Summarizing the regression results for Models 1-3, we can thus lend support to Hypotheses 1, 2 and 3, and accordingly for the main prediction of the Theory of Dyadic Morality.

In our design we measured harm and immorality perceptions along a continuum of six different payments in \$. The descriptive data analysis reveals that paying \$5 is perceived as neither harmful nor

	(1)	(2)	(3)	(4)	(5)
(Intercept)	1.13*** (0.08)	0.78*** (0.07)	0.49*** (0.07)	0.90*** (0.11)	1.08*** (0.14)
Global Harm	0.62*** (0.01)	0.40*** (0.02)	0.36*** (0.02)	0.35*** (0.02)	0.31*** (0.03)
Anger		0.39*** (0.02)	0.37*** (0.02)	0.37*** (0.02)	0.37*** (0.02)
Social Norm			0.12*** (0.01)	0.10*** (0.01)	0.10*** (0.01)
Payment \$1				-0.27*** (0.08)	-0.49*** (0.16)
Payment \$2				-0.33*** (0.08)	-0.57*** (0.16)
Payment \$3				-0.38*** (0.09)	-0.66*** (0.16)
Payment \$4				-0.45*** (0.09)	-0.70*** (0.15)
Payment \$5				-0.56*** (0.09)	-0.69*** (0.15)
Global Harm x Payment \$1					0.06 (0.04)
Global Harm x Payment \$2					0.07 (0.04)
Global Harm x Payment \$3					0.09 (0.05)
Global Harm x Payment \$4					0.08 (0.05)
Global Harm x Payment \$5					0.02 (0.04)
AIC	13324.98	12802.49	12739.39	12712.28	12715.72
BIC	13355.94	12839.64	12782.74	12780.39	12814.80
Log Likelihood	-6657.49	-6395.25	-6362.70	-6345.14	-6341.86
Num. obs.	3612	3612	3612	3612	3612
Num. groups: Payment nested within treatment	24	24	24		
Num. groups: Treatment	4	4	4	4	4
Var.: Payment nested within treatment (Intercept)	0.06	0.06	0.02		
Var.: Treatment (Intercept)	0.01	0.00	0.01	0.01	0.01
Var.: Residual	2.31	2.00	1.97	1.96	1.96
Marginal R^2	0.37	0.45	0.48	0.50	0.50
Conditional R^2	0.39	0.47	0.49	0.50	0.50

Notes: Regression results from mixed effects models: Immorality judgment is the dependent variable in all five models. Model 1 includes global harm perceptions as predictor, Model 2 adds anger as predictor and Model 3 adds social norm as predictor. The four framing treatments and the six payments in \$ nested within the framing treatments are included as random effects in Models 1-3. Model 4 and 5 are run to test whether global harm perceptions consistently predict immorality, irrespective of price paid. In Model 4, the six payments in \$ are included as predictor variables (factor variable). Model 5 adds an interaction between global harm perceptions and the six prices paid (factor variable). The four framing treatments are included as random effect in Models 4 and 5. Standard errors are provided in parentheses. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 3: Determinants of immorality judgment

	Df	AIC	BIC	logLik	deviance	Chisq	Chi Df	Pr(>Chisq)
Model 1	5	13324.98	13355.94	-6657.49	13314.98			
Model 2	6	12802.49	12839.64	-6395.25	12790.49	524.48	1	< .001
Model 3	7	12739.39	12782.74	-6362.70	12725.39	65.10	1	< .001
Model 4	11	12712.28	12780.39	-6345.14	12690.28	35.12	4	< .001
Model 5	16	12715.72	12814.80	-6341.86	12683.72	6.55	5	.256

Table 4: ANOVA for model comparison for testing the Theory of Dyadic Morality

immoral, whereas paying \$0 is perceived as both harmful and immoral. As a robustness check for the Theory of Dyadic Morality we thus explored whether harm perceptions predict immorality differently depending on the level of payment. We ran regressions (mixed effects models), including the six payments in \$ as (factor) predictor variable (Model 4, Table 3) as well as interactions between harm perceptions and the five positive payments (factor variable). We control for the between-subjects variance by including the framing treatments as random effect in both Models. Again, immorality is the dependent variable. The Theory of Dyadic Morality could be considered robust to differential payment levels if harm perception were consistently a significant predictor. Thus, we included the six payments as factor variable in the analysis. We find that payments of \$1, \$2, \$3, \$4, and \$5 are perceived as significantly less immoral than paying \$0 (Model 4). Thus, paying an amount of \$0 for consumption of online-news is perceived as most immoral compared to the five other payments in \$. This result is in line with the descriptive data analysis. We do not find a positive interaction between harm perceptions and the six payments in \$ (Model 5, see Table 3 for a summary of both Models), indicating that the predictive power of harm perceptions on immorality judgments may have been independent from any price paid.

As a robustness check for testing the TDM prediction, we tested whether the predictive power of harm perceptions, anger, and social norms differed between the four framing treatments (see Table 5).

	PWYW treatment	PWYC treatment	PWIIWTY treatment	PWYBIF treatment
(Intercept)	0.28** (0.10)	0.50*** (0.10)	0.85*** (0.15)	0.28** (0.09)
Global Harm	0.45*** (0.03)	0.43*** (0.03)	0.36*** (0.04)	0.19*** (0.03)
Anger	0.29*** (0.03)	0.35*** (0.03)	0.36*** (0.04)	0.48*** (0.03)
Social Norm	0.14*** (0.02)	0.05 (0.03)	0.07* (0.03)	0.25*** (0.03)
AIC	3153.71	3122.07	3343.90	3052.32
BIC	3182.72	3150.72	3372.76	3081.13
Log Likelihood	-1570.86	-1555.03	-1665.95	-1520.16
Num. obs.	930	876	906	900
Num. groups: Payment	6	6	6	6
Var.: Payment (Intercept)	0.00	0.00	0.04	0.00
Var.: Residual	1.71	2.04	2.29	1.72
Marginal R^2	0.54	0.45	0.41	0.58
Conditional R^2	0.54	0.45	0.42	0.58

Notes: Regression results from mixed effects models. Immorality judgment is the dependent variable in all four models, global harm perceptions, anger, and social norms are the predictors in all four models. Each model is computed separately for the four framing treatments and includes the six payments in \$ as random effect. Standard errors are provided in parentheses. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 5: Robustness of global harm perceptions, anger, and social norms predicting immorality by framing treatment

We ran four regressions (mixed effects models), one for each framing treatment, with immorality judgment as dependent variable, global harm perceptions, anger, and social norm as predictors, and prices paid as random effect. The results (see Table 5) suggest that global harm perceptions, anger,

and social norms predict immorality in all four framing treatments consistently, except for social norms in the PWYC treatment. For the latter treatment, social norm fails to reach statistical significance in prediction ($p = .11$). Global harm is the strongest predictor in all treatments, except for PWYBIF, where anger (coefficient = 0.48, $p < .001$) predicts immorality judgments stronger than perceived global harm (coefficient = 0.19, $p < .001$). Overall, model fit is best in the PWYBIF treatment ($AIC = 3052.32$, $R^2_{LMM(m)} = .58$, $R^2_{LMM(c)} = .58$).

5.2 Impact of political orientation

In order to do justice to the importance of political orientation in moral psychology, the following section focuses on the impact of political orientation on the process of immorality judgments. We ran regressions (mixed effects models) to detect differences in the process of immorality judgments with regard to political orientation. Table 6 summarizes the regression results.

	(1)	(2)	(3)
(Intercept)	0.49*** (0.07)	0.53*** (0.08)	0.29* (0.12)
Global Harm	0.36*** (0.02)	0.36*** (0.02)	0.44*** (0.04)
Anger	0.37*** (0.02)	0.37*** (0.02)	0.29*** (0.04)
Social Norm	0.12*** (0.01)	0.12*** (0.01)	0.19*** (0.03)
Political Orientation		-0.01 (0.01)	0.06* (0.03)
Global Harm * Political Orientation			-0.02* (0.01)
Anger * Political Orientation			0.02* (0.01)
Social Norm * Political Orientation			-0.02* (0.01)
AIC	12739.39	12740.41	12728.69
BIC	12782.74	12789.95	12796.80
Log Likelihood	-6362.70	-6362.21	-6353.34
Num. obs.	3612	3612	3612
Num. groups: Payment nested within treatment	24	24	24
Num. groups: Treatment	4	4	4
Var.: Payment nested within treatment (Intercept)	0.02	0.02	0.02
Var.: Treatment (Intercept)	0.01	0.01	0.01
Var.: Residual	1.97	1.97	1.96
Marginal R^2	0.48	0.48	0.49
Conditional R^2	0.49	0.49	0.49

Notes: Regression results from mixed effects models. Immorality judgment is the dependent variable in all three models, while global harm, anger, and social norms are the predictors in the three models. In Model 2 we add political orientation and in Model 3 we add interactions between political orientation and global harm, anger, and social norm, respectively. The six payments in \$ nested within the treatment are included as random effects in all three models. Standard errors are provided in parentheses. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 6: Political orientation and interaction with global harm, anger, and social norm as determinants of immorality judgment

In Model 1 we included harm perceptions, anger, and social norm as predictor variables, while adding the six prices paid nested within the treatment as random effects. Immorality judgment was the dependent variable. Model 2 was similar to Model 1 except that we added political orientation as predictor.

In Model 3 we added interactions between harm perceptions, anger, and social norm with political orientation, respectively.

Comparing Model 1 and Model 3 via an ANOVA, the fit indices of Model 1 ($AIC = 12739.39$, $BIC = 12782.74$, $R^2_{LMM(m)} = .48$, $R^2_{LMM(c)} = .49$) indicate a significantly better fit to the data ($\chi^2 = 18.71$, $p < .001$) than Model 3 ($AIC = 12728.69$, $BIC = 12796.80$, $R^2_{LMM(m)} = .49$, $R^2_{LMM(c)} = .49$). A significant main effect of political orientation on immorality judgment in Model 3 (coefficient = 0.06, $p = .038$) and significant interactions of political orientation with global harm (coefficient = -0.02, $p = .032$), anger (coefficient = 0.02, $p = .034$), and social norm (coefficient = -0.02, $p = .018$), were found. As higher levels of political orientation indicate rather conservative attitudes and vice versa, Model 3 suggests that the more conservative a participant is, the more they perceive a given situation as immoral and the stronger is the influence of anger on their immorality judgments. On the other hand, the more liberal a participant indicates themselves to be, the stronger is the influence of harm perceptions and social norms on their immorality judgments.

To compare the process between political orientation groups more easily, we ran regressions for Conservative and Liberal participants separately. We included global harm, anger, and social norm as predictor variables and added the six payments in \$ nested within the treatment as random effects. Immorality judgment was the dependent variable in both models. The estimates for each of the predictors vary substantially, as can be seen in Table 7.

	Conservative	Liberal
(Intercept)	0.53*** (0.09)	0.37*** (0.08)
Global Harm	0.29*** (0.03)	0.44*** (0.02)
Anger	0.47*** (0.03)	0.33*** (0.02)
Social Norm	0.06* (0.03)	0.13*** (0.02)
AIC	3814.04	5746.07
BIC	3848.81	5784.15
Log Likelihood	-1900.02	-2866.03
Num. obs.	1062	1704
Num. groups: Payment nested within treatment	24	24
Num. groups: Treatment	4	4
Var.: Payment nested within treatment (Intercept)	0.00	0.00
Var.: Treatment (Intercept)	0.00	0.01
Var.: Residual	2.10	1.69
Marginal R^2	0.45	0.57
Conditional R^2	0.45	0.57

Notes: Regression results from two mixed effects models. Immorality judgment is the dependent variable in both models, while global harm, anger, and social norms are the predictors in both models. The six prices paid nested within the treatment are included as random effects in both models. In Model 1 we run the regression for a subsample of Conservative subjects. In Model 2 we run the regression for a subsample of Liberal subjects. Standard errors are provided in parentheses. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 7: Determinants of immorality judgment split by Conservative and Liberal subjects

For Conservative participants, anger was the strongest predictor (coefficient = 0.47, $p < .001$), the estimate for global harm was smaller (coefficient = 0.29, $p < .001$), followed by social norm (coefficient = 0.06, $p = .014$). For Liberal subjects, global harm was the strongest predictor (coefficient = 0.44, $p < .001$), anger the second strongest (coefficient = 0.33, $p < .001$), followed by social norm (coefficient = 0.13, $p < .001$). Because of varying subsample sizes, fit indices of the regression models for Conservative subjects (AIC = 3814.04, BIC = 3848.81, $R^2_{\text{LMM(m)}} = .45$, $R^2_{\text{LMM(c)}} = .45$) and Liberal subjects (AIC = 5746.07, BIC = 5784.15, $R^2_{\text{LMM(m)}} = .57$, $R^2_{\text{LMM(c)}} = .57$) could not be tested for statistically significant differences.

Result 2 *The predictive power of global harm, anger, and social norms on immorality judgments is different for conservatives and liberals. Anger has a stronger predictive power on immorality judgments for conservatives, harm for liberals.*

Summarizing, the results suggest that immorality judgments might be construed differently for Liberal and Conservative participants, respectively. Thus, we suggest Hypothesis 4 should be rejected.

5.3 Impact of framing effects

The analysis regarding framing effects consists of two parts. First, we explore whether perceptions of immorality, harm perceptions, anger, and social norms differ between the four framing treatments (Hypothesis 5). Second, we investigate whether own payment intentions differ between the four framing treatments (Hypothesis 6).

To test Hypothesis 5, we analyze whether immorality judgments, harm perceptions, anger perceptions, and social norm perceptions differ between the four framing treatments and between the six payments in \$ (within each framing treatment). Table 8 provides summary statistics of immorality, global harm, anger, and social norm perceptions by framing treatment for all six prices paid. Figures A.1 and A.2 in the Appendix visualize these results.

Low payments are perceived as more immoral, socially inappropriate, and produce more anger as compared to high payments; this finding is similar in all four framing treatments. Taking a closer look at immorality judgments, PWIIWTY and PWYBIF seem to be more morally laden than the PWYW and PWYC framings. The immorality judgment values are always higher in the first two treatments as compared to the latter. Similarly, looking at social norm perceptions, payments in the PWIIWTY treatment result in higher perceptions of social inappropriateness than the other three framings. Thus, immorality and social norm perceptions seem to be volatile in terms of framing effects whereas this does not seem to be the case for anger and global harm perceptions.

Variable	PWYW	PWYC	PWIIWTY	PWYBIF
Immorality across all six prices paid	2.61	2.62	2.91	2.80
Immorality: Payment \$0	3.85	3.87	4.34	4.18
Immorality: Payment \$1	2.90	3.04	3.36	3.24
Immorality: Payment \$2	2.56	2.60	2.92	2.74
Immorality: Payment \$3	2.30	2.25	2.53	2.43
Immorality: Payment \$4	2.12	2.06	2.29	2.21
Immorality: Payment \$5	1.92	1.89	2.03	2.01
Global Harm across all six prices paid	2.59	2.63	2.58	2.57
Global Harm: Payment \$0	3.84	3.98	3.95	4.00
Global Harm: Payment \$1	2.95	3.21	3.06	3.05
Global Harm: Payment \$2	2.63	2.63	2.62	2.47
Global Harm: Payment \$3	2.23	2.20	2.18	2.09
Global Harm: Payment \$4	2.05	1.93	1.94	1.94
Global Harm: Payment \$5	1.85	1.81	1.73	1.84
Anger across all six prices paid	2.34	2.32	2.47	2.41
Anger: Payment \$0	2.91	3.29	3.47	3.31
Anger: Payment \$1	2.43	2.68	2.77	2.66
Anger: Payment \$2	2.19	2.28	2.44	2.29
Anger: Payment \$3	2.13	2.04	2.16	2.10
Anger: Payment \$4	2.17	1.87	2.02	2.06
Anger: Payment \$5	2.23	1.79	1.96	2.03
Social Norm across all six prices paid	3.38	3.47	3.69	3.45
Social Norm: Payment \$0	5.32	5.40	5.47	5.52
Social Norm: Payment \$1	3.97	4.24	4.44	4.30
Social Norm: Payment \$2	3.31	3.50	3.83	3.54
Social Norm: Payment \$3	2.83	2.92	3.16	2.83
Social Norm: Payment \$4	2.55	2.57	2.81	2.40
Social Norm: Payment \$5	2.29	2.19	2.42	2.08

Table 8: Perceptions of immorality, global harm, anger, and social norms for all six prices paid by framing treatment

We ran four regressions to further analyze the data. Table 9 summarizes the results from the regressions (mixed effects models). In Model 1, immorality was included as dependent variable (DV), predicted by the framing treatments (dummy variables) and across the six payments in \$ (included as random effect). In Model 2 the underlying framework was identical, except for the dependent variable, which was global harm. A similar structure was implemented in Model 3 and Model 4 for anger and social norm, respectively. PWYW was used as the reference framing treatment in each of the Models.

The regression results further confirm the descriptive results, namely, that perceptions of immorality and social norms differ between the four framing treatments, whereas harm perceptions and anger do not. Immorality perceptions in the PWIIWTY and in the PWYBIF treatment are statistically significantly higher than in the PWYW treatment. The regression result thus further confirms that payments are rated on higher immorality levels in the PWIIWTY and the PWYBIF treatments than in the PWYW treatment. There are no differences regarding global harm perceptions and anger between PWYW and the three other framing treatments. Social norm perceptions are statistically significantly higher in the PWIIWTY treatment compared to the PWYW treatment. This result provides suggestive evidence that payments are perceived as more socially inappropriate in the PWIIWTY treatment compared to the PWYW treatment. Whether payments are perceived as socially inappropriate or not thus seems to depend on the framing of the pricing mechanism.

	Model 1: Im- morality as DV	Model 2: Global Harm as DV	Model 3: Anger as DV	Model 4: Social Norm as DV
(Intercept)	2.61*** (0.32)	2.59*** (0.33)	2.34*** (0.20)	3.38*** (0.49)
Pay-What-You-Can	0.01 (0.09)	0.04 (0.08)	-0.02 (0.08)	0.09 (0.08)
Pay-What-It-Is-Worth-To-You	0.30*** (0.09)	-0.01 (0.08)	0.13 (0.08)	0.31*** (0.08)
Pay-What-You-Believe-Is-Fair	0.19* (0.09)	-0.02 (0.08)	0.07 (0.08)	0.07 (0.08)
AIC	14761.77	14370.61	14380.52	14399.06
BIC	14798.92	14407.76	14417.68	14436.21
Log Likelihood	-7374.88	-7179.30	-7184.26	-7193.53
Num. obs.	3612	3612	3612	3612
Num. groups: Payment	6	6	6	6
Var.: Payment (Intercept)	0.59	0.64	0.22	1.39
Var.: Residual	3.44	3.09	3.10	3.11

Notes: Regression results from four regressions (mixed effects models). The framing treatments are included as main predictor (dummy variables) and the six prices paid as random effect. The dependent variable varies in the four Models presented. Model 1 includes immorality as dependent variable (DV), whereas in Model 2 global harm is included as DV. Model 3 includes anger as DV and Model 4 includes social norm as DV. Standard errors are provided in parentheses. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 9: Framing treatment as determinant of immorality (Model 1), global harm (Model 2), anger (Model 3), and social norm (Model 4) perceptions

Result 3 *Voluntary payment markets are morally laden domains. Own payment intentions, immorality judgments, and perceptions of social norms however are slightly different between the four treatments.*

Summarizing the results, Hypothesis 5 can be confirmed partly: Framings bring about different perceptions of immorality and social norms, but not of harm perceptions and anger. Framing treatments lead to different immorality perceptions, especially in the PWIIWTY and in the PWYBIF treatment, in which immorality ratings are higher than in the PWYW treatment. Harm perceptions do not differ between PWYW and the other three framing treatments, also anger perceptions do not differ between PWYW and the other three framing treatments across the six payments in \$. Finally, payments in the PWIIWTY treatment are rated on higher social norm levels than in the PWYW treatment, indicating that social norms are more strongly activated in this setting. Nevertheless, participants agree that paying \$0 and low payments are perceived as more immoral and socially inappropriate than medium and high payments. This result is similar in all four framing treatments.

In Hypothesis 6 we predicted differences regarding own payment intentions between the four framing treatments.¹² The mean values of the own payment behavior intentions are statistically significantly different between the four framing treatments (Kruskal-Wallis test, $p = .005$). On average, participants in the PWYW treatment would pay \$2.98 ($SD = 3.53$) for consuming the weekly online-news of Today's Globe. This payment is only slightly smaller than the average in the PWYC treatment ($M = 2.99$, $SD = 2.38$). In the PWIIWTY treatment subjects would pay on average \$3.10 ($SD = 3.24$). The highest

¹²We do not claim that our results (payment intentions) should be interpreted as real payments as we are aware of the hypothetical elements in our study design. Nevertheless, we assume that the differences between the treatments are not systematically affected by this hypothetical character but rather due to our framing treatment variation.

amount would be paid in the PWYBIF treatment, with \$3.63 on average ($SD = 2.97$). A nonparametric Mann-Whitney test indicates that the difference between PWYW and PWYBIF is statistically significant ($p < .001$). Also, the payments between PWYC and PWYBIF differ statistically significantly from each other (Mann-Whitney test, $p = .03$). Furthermore, the difference between PWIIWTY and PWYBIF is statistically significant (Mann-Whitney test, $p = .006$). All other differences are not statistically significantly different from each other.

Summarizing, Hypothesis 6 is partially supported. Payments significantly differ between the four framing treatments: they were highest in the PWYBIF treatment, followed by payments in the PWIIWTY treatment, then the PWYC treatment and lowest in the PWYW treatment. However, the differences between the four framing treatments were not statistically significant for all comparisons. In conclusion, the payments in the PWYBIF treatment were statistically significantly higher than in all other three framing treatments.

6 Discussion

Is immorality predicted by perceptions of harm, anger, and social norms, as postulated by the Theory of Dyadic Morality (Schein and Gray, 2018)? And are voluntary payment markets perceived as moral domain? Our results provide empirical evidence for the TDM. Specifically, the three core elements proposed by Schein and Gray (2018) predict judgments of immorality (i.e., the perception of a social norm violation, a negative affect (anger), and harm ratings) all significantly and positively. Most importantly, perceptions of harm consistently predict perceptions of immorality (1) above and beyond the other two predictors, and (2) across varying prices paid. This sheds light on two complementary contributions of the current work.

First, we show that the moral psychological variables captured by the Theory of Dyadic Morality translate into actual judgments in a highly realistic setting. This is crucial, considering the heavy reliance of previous studies on bizarre scenarios that might not be indicative of the situations people experience in their day-to-day lives (see, for a discussion, Schein and Gray, 2018). Although such bizarre scenarios might be useful to rule out possible confounds of immorality perceptions, they may be considered low in external validity. Real world settings, such as voluntary payment markets, increase the external validity. At the same time, they represent an interesting and critical environment to test the TDM as it is uncertain which behavior is perceived as harmful, as social norm violation, as producing anger, and finally, as immoral.

Importantly, our data analysis supports the core predictions of the Theory of Dyadic Morality in this real-world setting. Across different framing options, our results demonstrate that norm perceptions

and negative affect (specifically anger), predict judgments of immorality. In other words, the more participants think that Individual A's behavior is socially inappropriate (i.e., violating a norm), and the more they are angered by it, the more they evaluate it as immoral. Furthermore, beyond these norm perceptions and negative affect, perceptions of harm consistently retain predictive value, indicating that the more participants think the act of Individual A is harmful, the more they deem it immoral. We are aware that voluntary payment contexts are settings in which the kind of harm is not physical but rather monetary (in cases where buyers pay nothing or low prices). Referring to Gray and Keeney (2015), we take into account that harm perceptions and immorality judgments might be lower than in scenarios where physical harm is investigated. Nevertheless, we see voluntary payment contexts as realistic examples in which immorality plays a role, in which harmful acts are possible, and in which the moral agent and the moral patient can be identified rather easily. Our data indeed shows moral typecasting processes, namely, that in voluntary payment markets buyers are perceived as moral agents and sellers as moral patients. To the best of our knowledge, these findings present a first direct test of the postulates of the Theory of Dyadic Morality in an applied everyday context, and lend great support to the theory.

Second, we demonstrate that voluntary payment markets are morally laden and that individuals seem to draw on their own cognitive and emotional compasses to form judgments of right and wrong in the absence of a posted price. The majority of subjects perceive payments of zero as very socially inappropriate and very immoral, whereas high payments are perceived as very socially appropriate and not immoral. Due to the ascending order in which the prices were presented we were not able to control for order effects. Whether the order of the prices paid might have served as reference point influencing the evaluations could serve as the basis for future research. Nevertheless, there is still a lot of moral wiggle room when it comes to immorality perceptions and social appropriateness for medium-high payments. Social norm perceptions seem to be especially volatile towards varying payments in \$. This opens the door for future research to further investigate this social component of voluntary payment settings and to explore how it can be utilized to optimize the setup for both buyers and sellers. In addition, future research on the underlying mechanisms of immorality perceptions in voluntary payment markets is necessary. It remains an open question, for instance, whether immorality perceptions in voluntary payment markets are shaped by buyers' reciprocity preferences in order to reward the trust of the sellers. Such a result would be in line with the assumption that markets can act as "moral teachers" (Storr and Choi, 2019, p. 195).

We also sought to investigate whether political orientation influences the process of immorality judgments as political orientation is an intensively debated aspect in moral psychology (see, e.g., Graham et al., 2009; Schein and Gray, 2018). While the moral typecasting process is similar for Liberal, Moderate,

and Conservative subjects, the process of immorality judgments is not. We explored whether liberals and conservatives had a general tendency to report higher perceptions of norm violations, anger, or harm perceptions. Our results suggest that they do not. Considering extensive literature documenting higher moralization among conservatives, this is a striking finding. However, when comparing the coefficients for the three core predictors we found differences in the magnitude of prediction. For liberals, the model works as suggested by Schein and Gray (2018): global harm is the strongest predictor, followed by anger and social norms, with all three being significant predictors. For conservatives, however, we find anger to be a stronger predictor than global harm, with social norms being the weakest predictor of immorality judgments. Given the similar overall anger perceptions between liberals and conservatives, this difference cannot be related to differences in the mean values. Future research can build on these results to further scrutinize the potential differences in moral decision making between liberals and conservatives.

Taken together, it may be that these results tie into the literature on domain specificity of political cognition. Contesting the assumption that liberals and conservatives differ fundamentally and systematically in their psychological make-up, several studies found differences were dependent on the context or domain rather than political orientation only. This applies *inter alia* to complex thinking (Conway et al., 2016), attitude formation and sensitivity to negative information (Fiagbenu et al., 2019), causal attributions (Morgan et al., 2010), intolerance toward outgroups (Brandt et al., 2014), or resistance to change (Proch et al., 2019). Similar to this strain of literature, the Theory of Dyadic Morality does not reject the fact that there may well be level differences in the predictive magnitude of the three components. This assumption implies that the mechanism itself is the same once all three components consistently predict the perception of immorality, regardless of their respective magnitudes. Applying insights from the literature on domain specificity for the present research, one could assume that either newspaper consumption or voluntary payment markets themselves may present a non-neutral and somewhat politically laden domain. Future research should thus vary the domain in order to get a more fine-grained idea of the underlying process and its interaction with political ideology.

Are the three core elements of the Theory of Dyadic Morality and immorality judgments as well as own payment intentions in voluntary payment markets sensitive towards framing effects? The results indicate that participants' own payment intentions differ significantly between the four framing treatments, being partly in line with previous research. Highest payment intentions occur in the PWYBIF treatment, while lowest payment intentions occur in the PWYW treatment. Payment intentions in the PWIIWTY and the PWYC treatment are situated in between. In line with previous findings (Sleesman and Conlon, 2017), payments are significantly higher in PWYBIF compared to PWYW. Being not fully in line with what was previously found (Cui and Wiggins, 2017; Machado and Sinha, 2015; Saccardo et al., 2021), we find that payments in PWIIWTY or in PWYC are not significantly higher than in PWYW.

These differences between our results and other studies might be explained in part by a different product used (online-news instead of cookies, entrance at a jazz club, drinks at a coffee shop, or snacks) or by collecting data online and not in the field or in the laboratory. The hypothetical character of our study design resulting in payment intentions instead of payments might also explain the differences. However, as we compare four framing variations in one study design, our results go beyond previous studies. We show that payment intentions in the PWYBIF treatment are significantly higher than in the PWYC treatment and are also significantly higher compared to PWIIWTY. Summarizing these results, the PWYBIF framing leads to higher voluntary payments than in the three other framings (PWYW, PWYC and PWIIWTY). It is thus crucial to be aware of the effects framing can have on voluntary payment intentions to the success of any practical application of voluntary payment mechanisms.

Regarding framing effects on the process of immorality judgments, our results suggest differences between the framing treatments for immorality and social norm perceptions, but not for anger and global harm perceptions. Our finding of framing effects is consistent with existing empirical evidence on framing effects. More specifically, the PWIIWTY and the PWYBIF treatments are perceived as more morally laden than the PWYW and the PWYC treatments. Regarding social norm perceptions, the results show that paying \$0 is perceived as most socially inappropriate and that paying \$5 is perceived as most socially appropriate, irrespective of the framing treatment. We find this pattern of social norm perceptions in all four framing treatments. Nevertheless, social norms seem to be more strongly activated in the PWIIWTY treatment compared to the PWYW treatment.

Our findings lend further support to the conclusion that the predictions of the Theory of Dyadic Morality are robust toward framing effects. Global harm, anger, and social norm perceptions consistently predict immorality judgments, except in the PWYC treatment. Particular to the PWYC treatment, we deem it interesting that social norm does not significantly predict immorality judgments. One explanation for this result might be our finding that low payments are also perceived as socially appropriate in the PWYC treatment. Hence, it seems to be the case that PWYC activates social norms less strongly than the other three framing treatments.

As predicted by the Theory of Dyadic Morality, global harm is the strongest predictor in all treatments, with the exception of PWYBIF. In the PWYBIF treatment, anger is the strongest predictor. We find a similar result in our regression model including global harm, anger, and social norms as determinants of immorality judgments: anger predicts immorality slightly more than global harm. We further find that model fit was best in the PWYBIF treatment, also similar to the regression model including global harm, anger, and social norms as determinants of immorality judgments. This result shows that besides global harm, anger seems to be an important predictor of immorality judgments, too, which should be reflected upon in the theoretical literature concerning immorality judgments.

This study comes with a few limitations that we believe are worth examining. First and foremost,

it should be noted that one main goal of this work is to specifically test the Theory of Dyadic Morality based on its own core claims. This means we tested the TDM and its prediction as if they were correct, exploring the usefulness of the TDM in predicting immorality judgments in voluntary payment settings. This in no way evaluates the TDM regarding its conceptual strengths and weaknesses, nor does it evaluate the relative merit(s) of the TDM in comparison to the numerous other competing conceptualizations of immorality judgments, including Moral Foundations Theory (Graham et al., 2013), Relationship Regulation Theory (Rai and Fiske, 2011), and others. Future research could directly test the divergent predictions of these different theories. In addition, future studies might also focus on contentious issues within the framework of TDM – for example, whether its core components (anger, harm, and social norm violation) have to coexist to produce immorality judgments (i.e., necessary conditions), or whether the existence of one or two of the components already leads to immorality judgments (i.e., sufficient conditions). These aspects have not been discussed to their full extent yet. Accordingly, it is not known to date how the three core components relate to each other if only one or two of them are at work, for instance in those cases where a social norm violation and harm are perceived without evidencing negative affect. This topic may be addressed in further studies. Finally, it is a question for future research to investigate the role of intention and the actual damage suffered in the process of immorality judgments. The TDM assumes that harm involves an intentional agent producing damage to a patient (Schein and Gray, 2018). However, recent studies (see, e.g., Hechler and Kessler, 2018) already show that the intention – without actually damaging somebody – leads to moral outrage and perceptions of immorality. Future studies could fruitfully explore the predictions of the TDM with and without intentionality of harm.

Second, we chose to test the TDM by relying on its authors' claims about the falsifiability of dyadic morality: "dyadic morality would be falsified if – when controlling for norms and affect – intuitive perceptions of harm are not causally linked to intuitive perception of immorality" (Schein and Gray, 2018, p. 27). This means that we took an additive, rather than an interactive approach to the predictors. One could potentially argue that a better interpretation of this constructivist theory would lead to the latter, meaning that one should expect – and therefore test – whether a three-way interaction between norms violations, harm perceptions, and negative affect predicts immorality judgments. Indeed, we ran such a somewhat stricter test of the TDM (three-way interaction) additional to our regression analysis focusing on the three core predictors as main effects. We found that while all three main effects significantly predict immorality judgments, two-way interactions are only significant between negative affect and norms violation as well as between social norm violations and global harm. The three-way interaction term also reaches significance, suggesting that the interaction between negative affect and norms violation may be different across different levels of harm. The more harm participants perceive, the weaker the interaction becomes. These results must be interpreted with caution as, consistent with our theorizing,

they were highly intercorrelated, perhaps hinting at the primacy of harm in the formation of immorality judgments and the potentially substitutional nature of negative affect and norms violation. Interpreted conservatively, these findings suggest that, if an action provokes harm perceptions to a lesser degree, then the interaction between negative affect and norm violations become more vital in predicting immorality judgments.

Furthermore, Schein and Gray's recommendation clearly involves a causal, rather than correlational claim. While we acknowledge that the correlational nature of our data limits our ability to make such a causal claim, we believe that an experimental design that attempts to exogenously vary harm perceptions includes an equal amount of limitations, albeit on different grounds. Indeed, one of the main tenets of TDM, and of most prominent psychological theories of moral judgments, is that "harm is in the eye of the beholder" (Schein and Gray, 2018, p. 8). We therefore argue that a study which claims to manipulate, rather than to measure harm, misses an essential insight of moral psychology – namely, that researchers ought not to impose their perceptions of harm on those of their participants. Similar argumentation can be applied in the case of anger and social norms. In light of this, we believe that a correlational design can help to shed light on these psychological processes without risking such top-down assumptions. Of course, further research could take an experimental approach and combine its findings with ours and others' correlational results to further illuminate the merits of the TDM and related theories of moral judgments. We believe that such a methodologically mixed approach is most likely to generate reliable knowledge for the field.

Additionally, we are aware of a number of limitations regarding the framings and product used in this study. One limitation is that our design is not exhaustive regarding the framings of voluntary payment contexts. We decided to explore the four most common and scientifically already investigated framings (Gerpott, 2017) Pay-What-You-Want, Pay-What-You-Can, Pay-What-It-Is-Worth-To-You, and Pay-What-You-Believe-Is-Fair. Thus, we excluded for instance semantically similar framings like Pay-What-You-Wish, Pay-What-You-Will, Pay-What-You-Like and framings which represent different mechanisms (focus on discount instead of paid price), such as Rebate-What-You-Want. Differences or similarities of these excluded framings to our four tested ones might be addressed in future research.

Another limitation might be seen in the hypothetical character of our study design, both regarding the payment intentions and the fictitious online-news as used product. As this feature of our experiment was randomized across treatments, we perceive the impact on treatment differences as low. Nevertheless, it remains an open question for future research whether our findings can be replicated in non-hypothetical settings (e.g. existing online-news websites) and for other products. Additional research with other products and settings is required, especially regarding the question of whether other types of voluntary payment markets (e.g. donations, tipping, etc.) are perceived as moral domain(s). Regarding the buying situation, it would be of interest to examine whether being observed during the payment decision matters

for immorality judgments. As known from research on social image concerns, voluntary payments can, for instance, be affected by making the payment publicly known to the seller or other buyers (Dorn and Suessmair, 2016, 2017; Gneezy et al., 2012; Hilbert and Suessmair, 2015; Hofmann et al., 2021; Regner and Riener, 2017; Schlüter and Vollan, 2015). Whether immorality and social norm perceptions also differ between anonymous and public voluntary payment settings remains unknown: any such insight(s) could, however, allow for a better understanding of the underlying mechanisms affecting immorality judgments in voluntary payment markets.

We are aware, furthermore, that Pay-What-You-Can is a special case compared to the other three framings. The relevant anchor for an individual's perception in the PWYC framing might not be the subjective evaluation of immorality so much as the available budget or income of the buyer. If a consumer's budget is limited, it might be socially appropriate to pay a lower amount, which would not be the case when the consumer has a bigger budget at their disposal. As our result indicates, social norms are no predictor of immorality judgments and might be perceived differently in the PWYC framing, which thus speaks in favor of such an assumption. Our study is limited by the fact that we have neither stated an amount nor kept the income of the buyer (Individual A) constant in our four treatments. If we had done so, this would have been an important confound variable for the subjective evaluations in the PWYW, PWIIWY, and PWYBIF framing treatments. However, it remains open for future research to further investigate immorality perceptions in a PWYC treatment by keeping the income of the consumer constant in order to rule it out as an affective variable.

Finally, future research on the role of social norms in differently framed voluntary payment contexts is desirable in order to extend the knowledge about the interplay of social norm perceptions and voluntary payment settings. Several other questions remain to be addressed in future research, including: differences in the process of immorality perceptions for repeat customers versus one-time visitors; the role of fairness perceptions in the process of immorality judgments; and how immorality perceptions occur in voluntary payment settings with reference prices provided. This latter aspect might reduce the uncertainty for a buyer about which price would be appropriate and an influence on immorality judgments accordingly. Additionally, more experiments will be needed to verify whether differently activated immorality and social norm perceptions due to framing effects are driven by changes in beliefs, changes in perceptions or changes in preferences.

7 Conclusion

Summing up, this paper empirically tests the theoretical predictions of the Theory of Dyadic Morality (Schein and Gray, 2018) in the applied setting of voluntary payment markets for the first time. We aimed to examine immorality judgments to determine if they could be predicted by the three core

elements of the Theory of Dyadic Morality (perceptions of global harm, anger, and social norms) as well as the extent to which voluntary payment markets are morally laden. Furthermore, we explored whether these measured immorality judgments, perceptions of harm, anger, and social norms as well as a participants' own payment intentions are affected by framing effects and whether they differ depending on the subjects' political orientation (which is a key variable in moral psychology research). In doing so, this paper therefore contributes to the previous literature by providing empirical evidence for Schein and Gray's Theory of Dyadic Morality in the applied setting of voluntary payments representing a critical environment for the theory.

We ran an online experiment using a between-subjects one-factorial design with four factor levels with 602 Americans being recruited via Amazon Mechanical Turk to complete the online experiment. The subjects were randomly assigned to one of the four voluntary payment settings, respectively framed as Pay-What-You-Want (PWYW), Pay-What-You-Can (PWYC), Pay-What-It-Is-Worth-To-You (PWII-WTY) and Pay-What-You-Believe-Is-Fair (PWYBIF). We then measured how voluntary payments for an online-news website were evaluated by the participants regarding their immorality, harm, anger, and social norm perceptions in addition to their own payment intentions.

Our empirical data lends support to the theoretical assumptions of the Theory of Dyadic Morality. We find that the theory's three core components do positively predict immorality judgments in the critical setting of voluntary payment markets: precisely the markets which capture everyday morality and in which it is unclear if perceptions of harm, anger, social norm violation, and immorality are activated. More specifically, we find that perceived harm consistently predicts immorality beyond social norm perceptions and anger. Differences in the process of immorality judgments are found with regard to political orientation where conservatives' immorality judgments were predict more strongly by anger than by harm. The results further show that most participants agree that paying \$0 for online-news is strongly socially inappropriate and also very immoral while high payments are evaluated as both morally and socially appropriate. This finding is similar in all four framing treatments. These results indicate that voluntary payment markets are indeed perceived as a moral domain and they support the relevance of social norms concerning the appropriateness of behavior in such markets. Our data further indicates that, in voluntary payment markets in which the price determination power is delegated to buyers, moral typecasting processes are also at work. Buyers are perceived as the moral agents who can generate harm and sellers are perceived as moral patients who can suffer from this harm. Furthermore, the data analysis reveals that both own payment intentions and the relevant measures of the Theory of Dyadic Morality in a voluntary payment context are sensitive to framing effects. In particular, PWIIWTY and PWYBIF are treatments which seem to be more morally laden than PWYW and PWYC. Finally, our research indicates that social norms are more strongly activated in the PWIIWTY framing whereas the PWYBIF framing leads to significantly highest voluntary payments intentions.

References

- Andreoni, James**, “Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments,” *The Quarterly Journal of Economics*, 1995, 110 (1), 1–21.
- Armstrong Soule, Catherine A. and Robert Madrigal**, “Anchors and norms in anonymous pay-what-you-want pricing contexts,” *Journal of Behavioral and Experimental Economics*, 2015, 57, 167–175.
- Bartling, Björn, Ernst Fehr, and Yagiz Özdemir**, “Does Market Interaction Erode Moral Values?,” *The Review of Economics and Statistics*, 2023, 105 (1), 226–235.
- Brandt, Mark J., Christine Reyna, John R. Chambers, Jarret T. Crawford, and Geoffrey Wetherell**, “The ideological-conflict hypothesis: Intolerance among both liberals and conservatives,” *Current Directions in Psychological Science*, 2014, 23 (1), 27–34.
- Brañas-Garza, Pablo**, “Promoting helping behavior with framing in dictator games,” *Journal of Economic Psychology*, 2007, 28 (4), 477 – 486.
- Brewer, Marilynn B. and Roderick M. Kramer**, “Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing,” *Journal of Personality and Social Psychology*, 1986, 50 (3), 543–549.
- Charness, Gary, Anya Samek, and Jeroen Van de Ven**, “What is considered deception in experimental economics?,” *Experimental Economics*, 2021, pp. 1573–6938.
- Conway, Lucian Gideon, Laura Janelle Gornick, Shannon C. Houck, Christopher Anderson, Jennifer Stockert, Diana Sessoms, and Kevin McCue**, “Are conservatives really more simple-minded than liberals? The domain specificity of complex thinking,” *Political Psychology*, 2016, 37 (6), 777–798.
- Cui, Annie Peng and Jennifer Wiggins**, “What You Ask Changes What I Pay: Framing Effects in Pay What You Want Pricing,” *Journal of Marketing Theory and Practice*, 2017, 25 (4), 323–339.
- Dana, Jason, Roberto Weber, and Jason Kuang**, “Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness,” *Economic Theory*, 02 2007, 33, 67–80.
- Deutsch, Morton**, “Equity, Equality, and Need: What Determines Which Value Will Be Used as the Basis of Distributive Justice?,” *Journal of Social Issues*, 1975, 31 (3), 137–149.
- Dorn, Tim and Augustin Suessmair**, “Is it really worth it? A test of pay-what-you-want pricing strategies in a German consumer behaviour context,” *Global Business and Economics Review*, 2016, 18 (1), 82–100.

- and —, “Determinants in Pay-What-You-Want Pricing Decisions – A Cross-Country Study,” *American Journal of Industrial and Business Management*, 2017, 7 (2), 115–142.
- Engel, Christoph and David G. Rand**, “What does “clean” really mean? The implicit framing of decontextualized experiments,” *Economics Letters*, 2014, 122 (3), 386 – 389.
- Falk, Armin and Nora Szech**, “Morals and Markets,” *Science*, 2013, 340 (6133), 707–711.
- Feldhaus, Christoph, Tassilo Sobotta, and Peter Werner**, “Norm Uncertainty and Voluntary Payments in the Field,” *Management Science*, 2018, 65 (4), 1855–1866.
- Fiagbenu, Michael Edem, Jutta Proch, and Thomas Kessler**, “Of deadly beans and risky stocks: Political ideology and attitude formation via exploration depend on the nature of the attitude stimuli,” *British Journal of Psychology*, 2019.
- Frisch, Deborah**, “Reasons for Framing Effects,” *Organizational Behavior and Human Decision Processes*, 1993, 54 (3), 399 – 429.
- Gerlach, Philipp and Bastian Jaeger**, “Another frame, another game? Explaining framing effects in economic games,” in Astrid Hopfensitz and Emiliano Lorini, eds., *Proceedings of norms, actions, games (NAG 2016)*, Toulouse: Institute for Advanced Studies, 2016, pp. 1–10.
- Gerpott, Torsten J.**, “Pay-What-You-Want pricing: An integrative review of the empirical research literature,” *Management Science Letters*, 2017, pp. 35–62.
- Gneezy, Ayelet, Uri Gneezy, Gerhard Riener, and Leif D. Nelson**, “Pay-what-you-want, identity, and self-signaling in markets,” *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109 (19), 7236–7240.
- , —, **Leif D. Nelson, and Amber Brown**, “Shared Social Responsibility: A Field Experiment in Pay-What-You-Want Pricing and Charitable Giving,” *Science*, 2010, 329, 325–327.
- Graham, Jesse, Jonathan Haidt, and Brian A. Nosek**, “Liberals and Conservatives Rely on Different Sets of Moral Foundations,” *Journal of Personality and Social Psychology*, 2009, 96 (5), 1029.
- , —, **Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto**, “Moral foundations theory: The pragmatic validity of moral pluralism,” in “Advances in experimental social psychology,” Vol. 47, Elsevier, 2013, pp. 55–130.
- Gray, Kurt and Chelsea Schein**, “No Absolutism Here: Harm Predicts Moral Judgment 30x Better Than Disgust-Commentary on Scott, Inbar, & Rozin (2016),” *Perspectives on Psychological Science*, 2016, 11 (3), 325–329.

- **and Daniel M. Wegner**, “Moral typecasting: Divergent perceptions of moral agents and moral patients,” *Journal of Personality and Social Psychology*, 2009, *96* (3), 505–520.
- **and Jesse Graham**, *Atlas of Moral Psychology*, Guilford Publications, 2019.
- **and Jonathan E. Keeney**, “Disconfirming Moral Foundations Theory on Its Own Terms,” *Social Psychological and Personality Science*, 2015, *6* (8), 874–877.
- , **Chelsea Schein**, **and Adrian F. Ward**, “The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering,” *Journal of Experimental Psychology: General*, 2014, *143* (4), 1600–1615.
- Greiff, Matthias and Henrik Egbert**, “The Pay-What-You-Want Game and Laboratory Experiments,” *MPRA Paper*, 2016, (75222).
- **and** — , “A Review of the Empirical Evidence on PWYW Pricing,” *Economic and Business Review*, 2018, *20* (2), 169–193.
- , — , **and Kreshnik Xhangolli**, “Pay What You Want – But Pay Enough! Information Asymmetries and PWYW Pricing,” *Management and Marketing*, 2013, *9*, 193–204.
- Haidt, Jonathan**, “The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment,” *Psychological Review*, 11 2001, *108*, 814–34.
- , *The Righteous Mind: Why Good People are Divided by Politics and Religion*, New York: Vintage, 2012.
- , **Fredrik Bjorklund**, **and Scott Murphy**, “Moral Dumbfounding: When Intuition Finds No Reason,” *Unpublished Manuscript, University of Virginia*, 2000, pp. 191–221.
- Haws, Kelly and William Bearden**, “Dynamic Pricing and Consumer Fairness Perceptions,” *Journal of Consumer Research*, 2006, *33* (3), 304–311.
- Hechler, Stefanie and Thomas Kessler**, “On the difference between moral outrage and empathic anger: Anger about wrongful deeds or harmful consequences,” *Journal of Experimental Social Psychology*, 05 2018, *76*, 270–282.
- Hilbert, Leon Paul and Augustin Suessmair**, “The Effects of Social Interaction and Social Norm Compliance in Pay-What-You-Want Situations,” *American Journal of Industrial and Business Management*, 2015, *5* (8), 548–556.
- Hoch, Stephen J. and Young-Won Ha**, “Consumer Learning: Advertising and the Ambiguity of Product Experience,” *Journal of Consumer Research*, 09 1986, *13* (2), 221–233.

- Hofmann, Elisa, Michael E. Fiagbenu, Asri Özgümüş, Amir M. Tahamtan, and Tobias Regner**, “Who is watching me? Disentangling audience and interpersonal closeness effects in a Pay-What-You-Want context,” *Journal of Behavioral and Experimental Economics*, 2021, *90*, 101631.
- Horberg, Elizabeth J., Christopher Oveis, and Dacher Keltner**, “Emotions as Moral Amplifiers: An Appraisal Tendency Approach to the Influences of Distinct Emotions upon Moral Judgment,” *Emotion Review*, 2011, *3* (3), 237–244.
- Inbar, Yoel, Paul Bloom, and David A. Pizarro**, “Disgusting Smells Cause Decreased Liking of Gay Men,” *Emotion*, 2012, *12* (1), 23–27.
- Jang, Hyunkyu and Wujin Chu**, “Are Consumers Acting Fairly Toward Companies? An Examination of Pay-What-You-Want Pricing,” *Journal of Macromarketing*, 2012, *32* (4), 348–360.
- Jost, John T. and Aaron C. Kay**, *Social Justice: History, Theory, and Research*, Vol. 2, Hoboken, NJ: Wiley,
- , **Jack Glaser, Arie W. Kruglanski, and Frank J. Sulloway**, “Political Conservatism as Motivated Social Cognition,” *Psychological Bulletin*, 2003, *129* (3), 339.
- Kahneman, Daniel and Amos Tversky**, “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 1979, *47* (2), 263–291.
- Kim, Ju-Young, Martin Natter, and Martin Spann**, “Pay What You Want: A New Participative Pricing Mechanism,” *Journal of Marketing*, 2009, *73* (1), 44–58.
- Kochanska, Grazyna, Katherine C. Coy, and Kelly Tobin Murray**, “The development of self-regulation in the first four years of life,” *Child Development*, 2001, *72*.
- Korenok, Oleg, Edward L. Millner, and Laura Razzolini**, “Taking, giving, and impure altruism in dictator games,” *Experimental Economics*, 2014, *17* (3), 488–500.
- Krämer, Florentin, Klaus M. Schmidt, Martin Spann, and Lucas Stich**, “Delegating pricing power to customers: Pay What You Want or Name Your Own Price?,” *Journal of Economic Behavior & Organization*, 2017, *136*, 125–140.
- Krupka, Erin L. and Roberto A. Weber**, “Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?,” *Journal of the European Economic Association*, 2013, *11* (3), 495–524.
- Kunter, Marcus**, “Exploring the Pay-What-You-Want payment motivation,” *Journal of Business Research*, 2015, *68* (11), 2347–2357.

- Liberman, Varda, Steven M. Samuels, and Lee Ross**, “The Name of the Game: Predictive Power of Reputations versus Situational Labels in Determining Prisoner’s Dilemma Game Moves,” *Personality and Social Psychology Bulletin*, 2004, 30 (9), 1175–1185.
- Machado, Fernando and Rajiv K. Sinha**, “The Viability of Pay-What-You-Want Pricing,” 2015. Working Paper.
- Mak, Vincent, Rami Zwick, Akshay R. Rao, and Jake A. Pattaratanakun**, “”Pay what you want” as threshold public good provision,” *Organizational Behavior and Human Decision Processes*, 2015, 127, 30–43.
- Marcus, Gary Fred**, *The Birth of the Mind: How a Tiny Number of Genes Creates The Complexities of Human Thought*, New York: Basic Civitas Books, 2004.
- Morgan, G. Scott, Elizabeth Mullen, and Linda J. Skitka**, “When values and attributions collide: Liberals’ and conservatives’ values motivate attributions for alleged misdeeds,” *Personality and Social Psychology Bulletin*, 2010, 36 (9), 1241–1254.
- Narwal, Preeti and Jogendra Kumar Nayak**, “How consumers respond to social norms: An evidence from pay-what-you-want (PWYW) pricing,” *Journal of Consumer Marketing*, 2019, 36 (4), 494–505.
- Natter, Martin and Katharina Kaufmann**, “Voluntary market payments: Underlying motives, success drivers and success potentials,” *Journal of Behavioral and Experimental Economics*, 2015, 57, 149–157.
- Proch, Jutta, Julia Elad-Strenger, and Thomas Kessler**, “Liberalism and conservatism, for a change! Rethinking the association between political orientation and relation to societal change,” *Political Psychology*, 2019, 40 (4), 877–903.
- Rai, Tage Shakti and Alan Page Fiske**, “Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality,” *Psychological Review*, 2011, 118 (1), 57–75.
- Rawls, John**, *A Theory of Justice*, Cambridge: Harvard University Press, 1971.
- Regner, Tobias and Gerhard Riener**, “Privacy Is Precious: On the Attempt to Lift Anonymity on the Internet to Increase Revenue,” *Journal of Economics & Management Strategy*, 2017, 26 (2), 318–336.
- **and Javier A. Barria**, “Do consumers pay voluntarily? The case of online music,” *Journal of Economic Behavior & Organization*, 2009, 71 (2), 395–406.
- Reno, Raymond R., Robert B. Cialdini, and Carl A. Kallgren**, “The transsituational influence of social norms,” *Journal of Personality and Social Psychology*, 1993, 64 (1), 104–112.

- Riener, Gerhard and Christian Traxler**, “Norms, moods, and free lunch: Longitudinal evidence on payments from a Pay-What-You-Want restaurant,” *The Journal of Socio-Economics*, 2012, 41 (4), 476–483.
- Royzman, Edward B., Kwanwoo Kim, and Robert F. Leeman**, “The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect,” *Judgment and Decision Making*, 2015, 10 (4), 296–313.
- Saccardo, Silvia, Charis X. Li, Anya Samek, and Ayelet Gneezy**, “Nudging generosity in consumer elective pricing,” *Organizational Behavior and Human Decision Processes*, 2021, 163, 91–104.
- Schein, Chelsea and Kurt Gray**, “The Unifying Moral Dyad: Liberals and Conservatives Share the Same Harm-Based Moral Template,” *Personality and Social Psychology Bulletin*, 2015, 41 (8), 1147–1163.
- **and** —, “Moralization and Harmification: The Dyadic Loop Explains How the Innocuous Becomes Harmful and Wrong,” *Psychological Inquiry*, 2016, 27 (1), 62–65.
- **and** —, “The Theory of Dyadic Morality: Reinventing Moral Judgment by Redefining Harm,” *Personality and Social Psychology Review*, 2018, 22 (1), 32–70.
- Schlüter, Achim and Björn Vollan**, “Flowers and an honour box: Evidence on framing effects,” *Journal of Behavioral and Experimental Economics*, 2015, 57, 186–199.
- Schmidt, Klaus M., Martin Spann, and Robert Zeithammer**, “Pay What You Want as a Market-ing Strategy in Monopolistic and Competitive Markets,” *Management Science*, 2015, 61 (6), 1217–1236.
- Sleesman, Dustin J. and Donald E. Conlon**, “Encouraging Prosocial Decisions: The Role of Fairness Salience and Uncertainty,” *Journal of Behavioral Decision Making*, 2017, 30 (2), 502–515.
- Sonnemans, Joep, Arthur Schram, and Theo Offerman**, “Public good provision and public bad prevention: The effect of framing,” *Journal of Economic Behavior and Organization*, 1998, 34 (1), 143 – 161.
- Storr, Virgil Henry and Ginny Seung Choi**, *Do Markets Corrupt Our Morals?*, Palgrave Macmillan, 2019.
- Tversky, Amos and Daniel Kahneman**, “The framing of decisions and the psychology of choice,” *Science*, 1981, 211 (4481), 453–458.
- **and** —, “Rational Choice and the Framing of Decisions,” *The Journal of Business*, 1986, 59 (4), 251–278.

Valdesolo, Piercarlo and David DeSteno, “Manipulations of Emotional Context Shape Moral Judgment,” *Psychological Science*, 2006, 17 (6), 476–477.

Vizuite-Luciano, Emili, Oktay Güzel, and José M. Merigó, “Bibliometric research of the Pay-What-You-Want Topic,” *Journal of Revenue and Pricing Management*, 2022, 17 (6), 1–14.

Watson, Amy, “Average price of digital newspaper subscriptions per week U.S. 2017 | Statista,” 2018.

Appendix

Additional material and further analyses

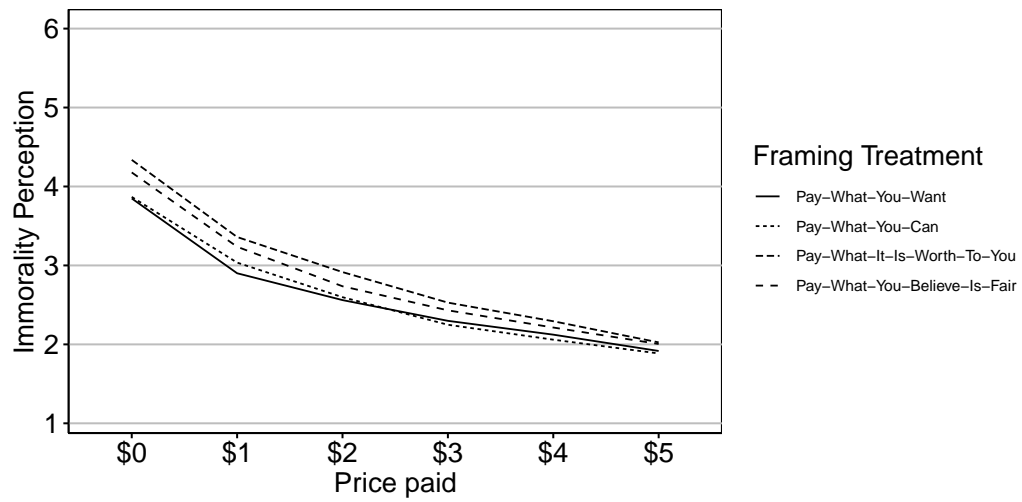
Variable	Mean	Std. Dev.	Min.	Max.	N
Immorality across all six prices paid	2.73	1.47	1	7	602
Immorality: Payment \$0	4.06	2.35	1	7	602
Immorality: Payment \$1	3.13	2.03	1	7	602
Immorality: Payment \$2	2.70	1.79	1	7	602
Immorality: Payment \$3	2.38	1.63	1	7	602
Immorality: Payment \$4	2.17	1.59	1	7	602
Immorality: Payment \$5	1.96	1.65	1	7	602
Global Harm across all six prices paid	2.59	1.35	1	7	602
Global Harm: Payment \$0	3.94	2.29	1	7	602
Global Harm: Payment \$1	3.06	1.96	1	7	602
Global Harm: Payment \$2	2.59	1.68	1	7	602
Global Harm: Payment \$3	2.18	1.47	1	7	602
Global Harm: Payment \$4	1.97	1.44	1	7	602
Global Harm: Payment \$5	1.81	1.53	1	7	602
Anger across all six prices paid	2.39	1.34	1	7	602
Anger: Payment \$0	3.24	2.28	1	7	602
Anger: Payment \$1	2.63	1.87	1	7	602
Anger: Payment \$2	2.30	1.59	1	7	602
Anger: Payment \$3	2.11	1.47	1	7	602
Anger: Payment \$4	2.03	1.51	1	7	602
Anger: Payment \$5	2.01	1.71	1	7	602
Social Norm across all six prices paid	3.50	1.22	1	7	602
Social Norm: Payment \$0	5.43	2.00	1	7	602
Social Norm: Payment \$1	4.23	1.92	1	7	602
Social Norm: Payment \$2	3.54	1.74	1	7	602
Social Norm: Payment \$3	2.94	1.54	1	7	602
Social Norm: Payment \$4	2.58	1.57	1	7	602
Social Norm: Payment \$5	2.25	1.78	1	7	602

Table A.1: Descriptive statistics on perceptions of immorality, global harm, anger, and social norms by paid price across treatments

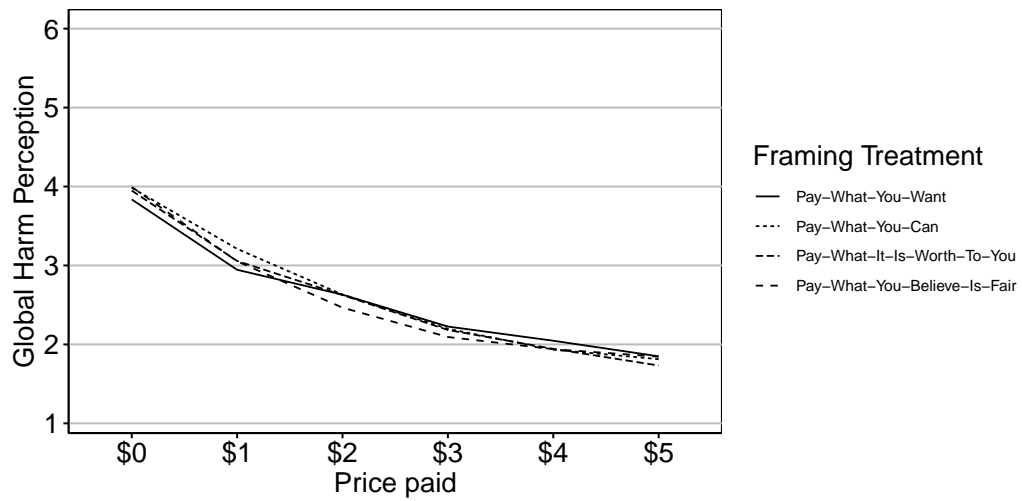
	Model 1: Immorality as DV	Model 2: Global Harm as DV	Model 3: Anger as DV	Model 4: So- cial Norm as DV
(Intercept)	4.10*** (0.10)	4.03*** (0.07)	3.20*** (0.07)	5.64*** (0.09)
Payment	-0.39*** (0.02)	-0.41*** (0.02)	-0.23*** (0.02)	-0.61*** (0.02)
AIC	14781.87	14387.00	14389.93	14443.55
BIC	14806.64	14411.77	14414.70	14468.32
Log Likelihood	-7386.94	-7189.50	-7190.97	-7217.78
Num. obs.	3612	3612	3612	3612
Num. groups: Treatment	4	4	4	4
Var.: Treatment (Intercept)	0.02	0.00	0.00	0.01
Var.: Residual	3.48	3.13	3.13	3.17
Marginal R^2	0.11	0.14	0.05	0.26
Conditional R^2	0.12	0.14	0.05	0.26

Notes: Regression results from four regressions (mixed effects models). Assuming equidistance between each level of the six payments in \$, we include payment as a continuous predictor and the four framing treatments as random effect. The dependent variable (DV) varies for each of the four Models presented. Model 1 includes immorality as DV, whereas in Model 2 global harm is included as DV. Model 3 includes anger as DV and Model 4 includes social norm as DV. * $p < .05$; ** $p < .01$; *** $p < .001$

Table A.2: Paid price as determinant of immorality (Model 1), global harm (Model 2), anger (Model 3), and social norm (Model 4) perceptions

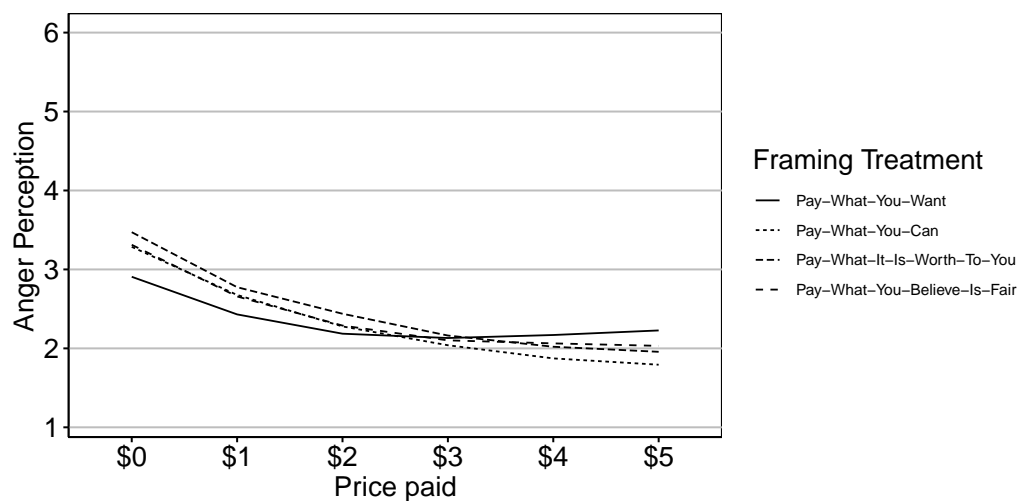


(a) Immorality perceptions for all six prices paid by framing treatment

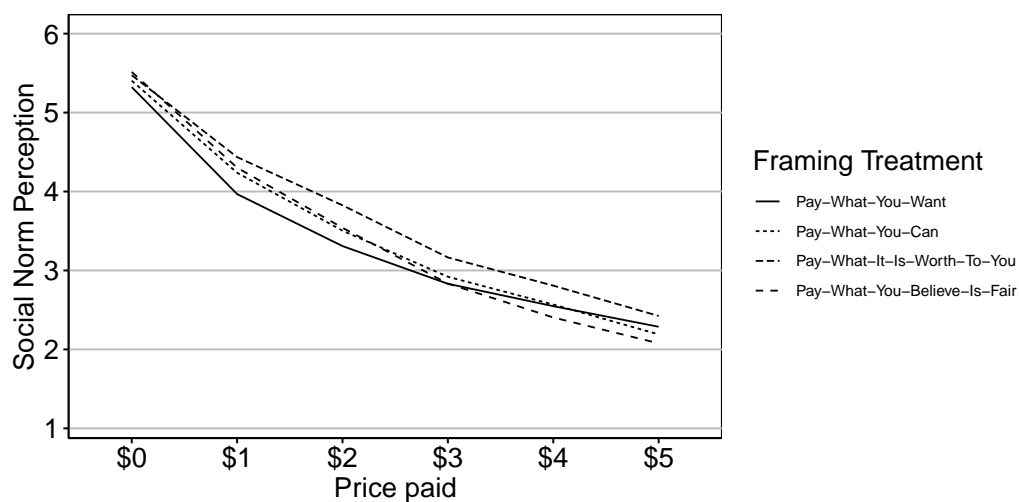


(b) Global harm perceptions for all six prices paid by framing treatment

Figure A.1: Immorality and global harm perceptions for all six prices paid by framing treatment



(a) Anger perceptions for all six prices paid by framing treatment



(b) Social norm perceptions for all six prices paid by framing treatment

Figure A.2: Anger and social norm perceptions for all six prices paid by framing treatment

IMPRESSUM

Jena Economics Research Papers

ISSN 1864-7057

Friedrich Schiller University Jena

Faculty of Economics and Business Administration

Carl-Zeiss-Str. 3

D-07743 Jena, Germany

Email: office.jerp@uni-jena.de

Editor: Silke Übelmesser

Website: www.wiwi.uni-jena.de/en/jerp

© by the author