Guilt Induced Self-Punishment as a Sign of Remorse
Rob M. A. Nelissen
Social Psychological and Personality Science published online 9 June 2011
DOI: 10.1177/1948550611411520

The online version of this article can be found at:
http://spp.sagepub.com/content/early/2011/06/09/1948550611411520

Published by:
SAGE
http://www.sagepublications.com

On behalf of:
Society for Personality and Social Psychology
Association for Research in Personality
European Association of Social Psychology
Society of Experimental and Social Psychology

Additional services and information for Social Psychological and Personality Science can be found at:

Email Alerts: http://spp.sagepub.com/cgi/alerts
Subscriptions: http://spp.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Guilt-Induced Self-Punishment as a Sign of Remorse

Rob M. A. Nelissen

Abstract
Why do people engage in self-punishment when they feel guilty? This article aims to bridge discrepant views that portray guilt either as an adaptive social emotion that is vital to the maintenance of social relations or as a maladaptive emotion that produces a host of negative self-directed responses. An experiment investigating the impact of various audience conditions on self-punishment tendencies suggested that even the negative self-directed responses that characterize certain episodes of guilt may originally serve an adaptive social function by acting as signals of remorse.

Keywords
self-punishment, guilt, prosocial behavior, psychopathology

Ritualized self-punishment is a universal aspect of human cultures throughout history. It seems to derive from a common belief that suffering brings relief from sin (Glucklich, 2001). The use of self-punishment to relieve guilt underlies the Roman Catholic principle of penitence and is also practiced in other religious traditions, such as on the Muslim Day of Ashura. Supporting that belief, recent findings established that self-punishment may be a behavioral response to guilt (Nelissen & Zeelenberg, 2009) that is indeed effective in reducing guilt feelings (Bastian, Jetten, & Fasoli, 2011). The guilt-induced tendency for self-punishment that arises if there are no possibilities to compensate the victim directly is called the Dobby effect (cf. Nelissen & Zeelenberg, 2009).

Still, the existence of a Dobby effect in itself does not explain why self-punishment purifies and/or alleviates guilt feelings. To the extent that self-punishment is a sign of true guilt, people may be more likely to be forgiven for their transgressions after displays of penitence. This would explain the guilt-alleviating effect of self-punishment, and would be in-line with views of guilt as an adaptive emotion that guards interpersonal relationships (Tangney, Stuewig, & Mashek, 2007). In the present article, we will test this relation-affirmation function of guilt-induced self-punishment.

On the other hand, guilt-induced self-punishment is also a symptom of various psychopathologies, and as such may be completely dysfunctional (Freud, 1916/1957). The discrepant perspectives on the functional nature of guilt-induced self-punishment probably originates from a distinction in research focus on the interpersonal or intrapersonal effects of guilt. The idea that guilt-induced self-punishment may function as a social signal is based on research on the interpersonal consequences of guilt, which shows that guilt instigates compensatory actions that restore damage to interpersonal relations (e.g., Baumeister, Stillwell, & Heatherton, 1994; De Hooge, Zeelenberg, & Breugelmans, 2007; Frank, 1988; Ketelaar & Au, 2003; Nelissen, Dijker, & De Vries, 2007; Tangney & Dearing, 2002). However, research on the intrapersonal consequences of guilt (e.g., Kugler & Jones, 1992; Lindsay-Hartz, de Rivera, & Mascolo, 1995) challenges this positive view by showing that feelings of guilt often result in destructive responses, which may underlie pathological conditions such as depression, obsessive compulsiveness, and psychosis (Fedewa, Burns, & Gomez, 2005; Harder, Cuttler, & Rockart, 1992; Lewis, 1971; Zahn-Waxler & Kochanska, 1990). By revealing that guilt-induced self-punishment is indeed a social signal whose objective is to affirm the relation between offender and victim, the present findings suggest that even the negative, self-directed responses to guilt at the intrapersonal level may originate from the adaptive interperson- nal functions of guilt.

The Objectives of Guilt-Induced Self-Punishment
In addition to the possibility that guilt-induced self-punishment is a social signal of remorse, the present research tested two
alternative hypotheses of the nature of negative self-directed responses to guilt: First, people may engage in self-punishment when feeling guilty as an emotion-regulation strategy aimed at self-affirmation. This would be consistent with self-affirmation theory (Steele, 1988), which holds that people are inclined to protect and restore their self-integrity after moral or interpersonal transgressions. According to this view, people engage in (ritual) behaviors to psychologically cleanse themselves (Zhong & Liljenquist, 2006). A self-affirmation objective of self-punishment would also be in-line with the view of guilt as a self-conscious emotion, elicited by negative appraisals of personal conduct (e.g., Lewis, 1971).

Secondly, guilt-induced self-punishment could also be a means to affirm one’s social value. This follows from the universal need to belong to a group and the fear of social exclusion (Baumeister & Leary, 1995). According to this view, social transgressions not only jeopardize the relation with a specific person but also one’s status as a group member. Therefore, after committing a transgression, people experience a need to avoid exclusion by showing others (but not necessarily the victim) that they possess desirable characteristics, and thus affirm their social value as group members (Caplovitz-Barrett, 1995). A social-affirmation objective of self-punishment would be in-line with the view of guilt as a social emotion, elicited through appraisals of others’ evaluation of the self (e.g., Kugler & Jones, 1992) but maybe more readily perceived as a behavioral response to shame rather than to guilt (see discussion).

Thirdly, the objective of guilt-induced self-punishment could be to affirm the relationship with the victim of one’s transgression, as predicted. This would be consistent with the proposed relationship restoring function of guilt (Baumeister et al., 1994; Tangney et al., 2007). According to this view, guilt-induced self-punishment is another way to restore the relationship with the victim (e.g., De Hooge et al., 2007; Ketelaar & Au, 1993; Nelissen et al., 2007). A relationship-affirmation objective would also be in-line with the view of guilt as a moral emotion, elicited by appraisals that signal violation of other peoples’ concerns (Haidt, 2003). We favor this objective, because it is the only one that acknowledges that feelings of guilt are primarily experienced in dyadic relationships and are not generally experienced in private nor expressed publicly to an audience (Baumeister et al., 1994).

The rationale behind this approach was that goal-directed behavior should be performed with greater intensity if the circumstances are more conducive to the attainment of the goal (Gollwitzer & Moskowitz, 1996). So, we predicted that people who felt guilty (but not in a control condition) would be more inclined to engage in self-punishment in the presence of their victim but not in the presence of a general audience or in the absence of an audience.

We only used female participants because the confederate in the audience conditions was also a female. We wanted to exclude confounding by male participants trying to show-off by engaging in more self-punishment. Our measure of self-punishment (ostensibly) involved self-inflicted pain through self-administered electrical shocks. Male participants might use pain-tolerance as a means to impress the confederate (e.g., Roney, Mahler, & Maestripieri, 2003).

Method

Participants and Design

A total of 85 female undergraduate students, $M_{age} = 20.0, SD = 4.0$, were randomly assigned to the experimental conditions and were seated in individual cubicles. The experiment consisted of two tasks, a guilt-induction procedure and a quiz that was used to collect the dependent measure of self-punishment. The study had a two (emotion: control vs. guilt) × three (audience: no vs. general vs. target) between subjects design. The willingness to self-administer electrical shocks was used as a measure for self-punishment.

Guilt-Induction

Participants played two rounds of a performance task, ostensibly with another participant. In this task, which has been demonstrated to be an effective guilt-induction manipulation (De Hooge, Nelissen, Breugelmans, & Zeelenberg, 2011), letters appeared rapidly on the screen in either a red or green color. To earn points, participants had to respond to green letters before they disappeared from the screen by pressing the corresponding letter on the keyboard. If the letter was red, the other player could earn points in a similar fashion. In the first round participants could earn lottery tickets for themselves, in the second round they earned lottery tickets for the other player. Importantly, both players needed to do well in order to reach the minimum level of 100 points to get the tickets.

After playing the first round of the letter task, all participants received bogus feedback that they earned the tickets due to good performance of both themselves and the other player. Participants subsequently played the second round of the letter task to earn tickets for the other player. In the guilt condition, the other player did not receive any tickets due to the participant’s bad performance. In the control condition, the other player also received the tickets.
Emotion Measures

Participants indicated to what extent they felt guilty, bad for what they did, and regretted their effort ($\alpha = .84$) on a scale from 1 (not at all) to 7 (very much). To verify that the observed effects are unique to feelings of guilt, and not driven by negative feelings in general, we also assessed to what extent participants experienced shame, a closely related but different social emotion. So, participants also indicated to what extent they felt ashamed, embarrassed, and bad about themselves ($\alpha = .82$).

The Quiz Procedure

Next, participants were escorted to another room and seated at a table. In the no-audience condition, participants were seated alone. In the audience conditions, a female confederate was seated at the opposite end of the table. The experimenter either mentioned explicitly that the confederate was not the same person with whom they were paired in the previous task (general audience condition) or that they were also paired to this person in the previous task (target audience condition). The experimenter then asked the confederate to put on the headphones, thus preventing further interaction.

A game of trivial pursuit was laid out on the table, and what appeared to be an electrical device was placed next to it. This device consisted of the contents of a digital video disc (DVD) player, in a wooden case, topped by a transparent cover. When seated, the participants received a sheet with the following instructions (in the audience conditions, the confederate was handed the same sheet of paper, which they started to read immediately, thus excluding nonverbal interaction as well):

You will be playing a quiz. For incorrect answers, you will have to administer yourself a mild shock. That is what the device on the table is for. With this task we examine if aversive stimuli can improve performance in knowledge tasks.

In the audience conditions, it was stated that participants were randomly selected to operate the shocking device, to justify them being the one administering the shocks and not the other person. The instructions continued:

The device is not plugged in yet, but will be as soon as we start the quiz. You first need to determine the shock level. You can do so by turning the knob. The red area indicates the highest, and the green area indicates the lowest shock level. Don’t worry; shocks will not be harmful, only discomforting. Still, the higher the shock levels, the more useful the results will be.

The final sentence was added to stress the (social) value of higher shock levels, thus making them potentially effective not only as signals of remorse but also as means of self- and social-affirmation. So, even in the absence of the target of guilt feelings, participants were explicitly offered a way (i.e., setting higher shocks) that helped them to establish their personal or social value by answering to the need of the experimenter.

After reading the instructions, participants turned the dial to set the shock level. After that, they called the experimenter back into the room.

Self-Punishment

The color of the area in which the participant placed the pointer on the knob allowed us to measure (willingness for) self-punishment on a 6-point scale.

Debriefing

After the experimenter recorded the shock level, participants were explained that the quiz would not be played and they were escorted back to their cubicles for individual debriefing. To measure aversion for the shocks, we asked participants to indicate on a scale from 1 (not at all) to 7 (very much) to what extent they were afraid to get electric shocks and thought the shocks would be painful ($r = .78$). To verify the effectiveness of the audience manipulations, we asked participants if they thought they were going to play the quiz alone or with someone else, and in case of the latter, if they had been paired with this person before or not. Next, we used a funneled procedure to probe for suspicion concerning the goal of the experiment, and the authenticity of the tasks, and the interactions.

Results

Manipulation Checks

None of the participants reported suspicion about the authenticity of the performance task scores and the interaction. All participants correctly identified their audience conditions in the quiz procedure. One participant recognized the shocking device as a broken DVD player and was excluded from further analysis. An emotion $\times$ audience analysis of variance (ANOVA) did not reveal differences in reported aversion for electric shocks between experimental conditions (all $F$s < 1). So, differences in self-punishment levels cannot be attributed to between-group differences in perceived aversive consequences of administering shocks.

The guilt-induction was successful. An emotion $\times$ audience ANOVA on differences in experienced guilt after the performance task revealed only a main effect of emotion, $F(1, 84) = 69.15, p < .001$, but not of audience, $F(1, 84) = 0.46, p = .635$, or of the interaction, $F(1, 84) = 0.08, p = .924$. Participants in the guilt no-audience ($M = 3.48, SD = 1.22$), the guilt general audience ($M = 3.38, SD = 1.37$), and the guilt target audience ($M = 3.53, SD = 1.10$) conditions, reported higher levels of experienced guilt than participants in the control no-audience ($M = 1.63, SD = 0.63$), the control general audience ($M = 1.38, SD = 0.62$), and the control target audience ($M = 1.76, SD = 0.91$) conditions. Feelings of shame were not affected by the emotion, $F(1, 84) = 1.33, p = .253$, nor by the audience manipulation, $F(1, 84) = .034, p = .967$, and we did not observe an interaction effect either, $F(1, 84) = 1.59, p = .210$. This indicated that the observed
punishment as the only possible signal of remorse. According to equity theory (Walster, Berscheid, & Walster, 1973), people will respond to their guilt feelings in a cost-efficient manner, meaning that they will prefer to repair their transgression with minimal costs to the self. As compensating the victim is often a less costly means to restore the inequity that resulted from the transgression than hurting the self (cf. Walster et al., 1973), prosocial behavior is the common reaction to feelings of guilt. Self-punishment then, is only a measure of last resort.

Cost-efficiency may also explain the absence of higher levels of self-punishment in all guilt conditions. Although this is admittedly speculative, it could be argued that the presently adopted measure of self-punishment (electrical shocks) was of a highly aversive nature. Consistent with the equity perspective on behavioral reactions to guilt, a response as costly as administering oneself electrical shocks is only cost-efficient if it stands a reasonable chance at attaining its desired goal. Apparently, participants were only willing to increase the shock level in circumstances that were conducive to the objective underlying their self-punishment tendency. As such, these results presented a robust indication that guilt-induced self-punishment is a signal of remorse and are the first to provide empirical support to the functionality of self-punishment behavior.

To some extent, the present findings exclude some alternative explanations as to why guilt elicits self-punishment. If self-affirmation were the objective of self-punishment (e.g., Zhong & Liljenquist, 2006), attainment of the objective would not depend on the type of audience, as this objective is strictly intrapersonal. Self-punishment would therefore be higher in all guilt conditions compared to the control conditions. If social-affirmation was the objective of self-punishment (cf. Caplovitz-Barrett, 1995), participants would have been more inclined to punish themselves in the presence of any audience rather than in private, no matter if the audience included their victim or not. As self-punishment was only higher in the target audience condition, however, it seems that relation-affirmation was the main objective of self-punishment. Any situation that did not include the victim did not enable the attainment of that objective.

Obviously, the present research has its limitations. Firstly, the experiment relied on a single manipulation of guilt. We are nevertheless confident that the present findings are more readily attributable to guilt than to other social emotions, particularly shame. We acknowledge that it is hard to completely exclude a role for shame, even with the type of manipulation check that we used, as people tend to confuse the labels shame and guilt. Feelings of shame and guilt tend to co-occur. Still, one key distinction is that the elicitation of guilt is almost exclusively related to transgressions toward others, whereas shame can also be experienced following general displays of incompetence (Smith, Webster, Parrott, & Eyre, 2002). Our manipulation of guilt was not based on incompetence as the participant had already performed better in the first round of the performance task. So, the result in the second round was more a lapse than a sign of incompetence. Therefore, the most salient aspect of the negative result in Round 2 was the...
disadvantage to the other participant, especially given the way the participants themselves had benefited from the other players’ effort. In addition, we believe that if shame would have contributed to people’s self-punishment tendencies, we would also have observed increased levels of self-punishment in the general audience condition. Shame-induced behavior is functionally targeted at reaffirming one’s social value (e.g., De Hooge et al., 2007; Haidt, 2003). Therefore, if shame played a role in the present findings, it would have instigated self-punishment in front of a general audience. (Remember that setting a higher shock level was presented as a socially valuable strategy, thus identifying oneself as a socially valuable person.) We believe that the absence of an increase in self-reported shame, the nature of the guilt-induction manipulation, and the lack of self-punishment in the general audience condition, renders an interpretation of our results as guilt-induced effects more likely than one in terms of shame-instigated action.

Still, to positively exclude alternative interpretations and to gain more support for the idea that self-punishment functions as a signal of remorse, future research should include different ways of manipulating audiences. Specifically, the present findings could be aligned with research on social-affiliation seeking (e.g., Kulik & Mahler, 2000; Schachter, 1959) to investigate if participants would actually seek out the targets of their guilt feelings or would rather seek the company of others or even prefer to stay alone. In addition, future research could also investigate whether the guilt-induced tendency for self-punishment may be functional in the sense that it enhances the effectiveness of apologies to elicit forgiveness (Hareli & Eisikovits, 2006). The present findings suggest that if an apology involves a certain amount of self-punishment it will be more effective, because self-punishment reliably signals remorse. In that respect, it would also be interesting to see if people consciously use self-punishment to elicit forgiveness in their victims. These and other strategies may provide further support to the proposed adaptive function of guilt-induced self-punishment.

A final limitation is that our findings rely on a single type of self-punishment (pain). We found it is hard to think of a negative, self-directed behavior that is more aversive than pain. Therefore, the anticipation of pain presented a very conservative test of our hypotheses. Still, the present findings should be extended in future research to test the generalizability of the present findings, if only to see whether the effect also holds for male participants as well.

Keeping these considerations in mind, we conclude that even though negative self-directed reactions to guilt may seem dysfunctional at first, they may in fact play a functional role in the interpersonal domain that is consistent with the general adaptive value of guilt as a guard of interpersonal relations. This does not preclude, however, that when displayed excessively, these same responses might be at the core of guilt-related psychopathologies. As such, the present findings bridge discrepant views on the nature of guilt.

Acknowledgment
I would like to thank Niels van de Ven and Yoel Inbar for their useful comments to earlier drafts of this manuscript.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


**Bio**

Rob M. A. Nelissen is an assistant professor at the Department of Social Psychology at Tilburg University, the Netherlands. His research focuses on the functional role of emotions in social behavior.