Causality and Blame Judgments of Negative Side Effects of Actions May Differ for Different Institutional Domains

Inna F. Deviatko1,2 and Kirill A. Gavrilov1,2

Abstract
Cognitive factors are known to influence lay assessments of causality and blame for negative side effects of intentional actions but specific social determinants of such assessments remain relatively unexplored. In a full-factorial, intraindividual experiment using two blocks of analogous vignettes constructed for two particular institutional action domains (“medical” and “corporate dress code”), we tested the propositions that causality and blame judgments differ between (a) domains and depend on (b) the type of action originator; (c) the type of damage; and (d) the “remoteness” of damage from the originator. Our data demonstrate a significant difference between two institutional action domains: actors in “medical”-related vignettes are generally estimated to be more causally effective and blameworthy than actors in “dress code”-related vignettes. In addition to the pronounced main effects of institutional domain as a factor influencing cause and blame judgments, we revealed few significant interaction effects of the latter with other experimental factors used for vignettes construction.

Keywords
cause, blame, attribution, type of damage, institutional domain of action, folk sociology

Introduction
Evaluation of accidental damage from an intentional legitimate action has been the subject of long debate among law scholars (Finkelstein, 2005; Hart, 1968; J. W. Turner, 1936). They usually consider the concept of criminal liability for negligence as a part of “a more general theory of responsibility for action” and assumed that “the criminal law is simply another form of ordinary, non-criminal responsibility” (Finkelstein, 2005, p. 581). These processes of unintended side effects evaluations made by lay people have been widely discussed in recent publications. Key issues are the relationship between ordinary people’s judgments of causality, responsibility, and blame, on one hand, and, on the other hand, the “folk psychology” as a system of cognitive processes and naive theories of mind and behavior (see e.g., Alicia, 2000; Cushman, 2008; Guglielmo & Malle, 2010; Knobe, 2003). Studies of folk psychology effects led researchers to believe that they have central importance in shaping moral judgments (Guglielmo et al., 2009).

Indeed, blame and responsibility attribution in everyday reasoning occurs within a wider context of beliefs and judgments about consciousness, goal directedness, supposed desires, reasons, and causal determinants of agents’ deeds. But there is another important source of blame, responsibility, and even intentionality judgments—“folk sociology.” This is a system of descriptive, normative, and explanatory knowledge of the social world secured by relevant cognitive processes. Folk sociology supports the possibility of deciding exactly which social norms have been broken in a specific situation with harmful effects for an agent himself or for other people; which social actions could be considered right or legitimate in this context; who is an actor, and so on.

In this article, we proceed as follows: in section “Background,” we analyze the theoretical background for our empirical research basing on previous theoretical work and empirical findings concerning our main experimental factors influencing lay judgment of cause and blame, namely, institutional domain and proximity of an initial action to the final negative outcome (1.1.1), type of social actor (1.1.2), and type of damage (1.1.3), concluding this with a general outline of our experimental design against the background of previous research (1.2). Then we describe in

1National Research University Higher School of Economics, Moscow, Russian Federation
2Institute of Sociology of FCTAS RAS, Moscow, Russian Federation

Corresponding Author:
Inna F. Deviatko, Sociology Department, National Research University Higher School of Economics, 11 Myasnitskaya ul., Room 333, Moscow 101000, Russian Federation.
Email: deviatko@gmail.com
detail the method (section “Method”) and our experimental results (section “Results”) which are discussed and summarized in sections “Discussion” and “Conclusions.”

Background

Folk psychology versus folk sociology in judgments of blame and cause: Institutional domains. As it has been shown before (Cova et al., 2016; Guglielmo & Malle, 2010; Knobe, 2003; see also Knobe et al., 2012), a perceiver’s judgments concerning the intentionality of the side effects of intentional actions depend not only on his or her evaluations of the negative and positive outcomes of the action and the actor’s conscious intentions but also on the perceiver’s view of the subjective desirability of expected externalities for the actor and actor’s ability to control them.

We assume that the very possibility of making such intentionality and blame inferences from an agent’s observed behavior also depends on socially distributed rights to act and obligations to control the negative side effects of one’s actions. These obligations have a rather limited range, depending on a specific action domain and on the possibilities of control from the actor’s side. No one is expected to control all of distant consequences of his or her actions. For example, a scientist who designed a cheap and highly intelligent robot is hardly accountable for a decrease in the job market demand for a certain category of qualified workers as the result of his invention. It is important that this is not a question of foreseeability because an actor, at least in principle and not facing restrictions related to time or mental ability, could expect almost all of these negative side effects. Rather, it appears that actors are often considered by others as exempt from obligations to prevent such “remote” damage even in a case of non-negligible likelihood of the negative outcome.

These limitations on an actor’s responsibility for intelligently expected, most likely unintended damage can be considered as implied by the very concept of a legitimate individual action. We imply that a person has the right to perform an action that he or she is interested in and allowed to control without granting someone potentially affected by this action the right to control it. (Compare this with the treatment of norms by J. S. Coleman, who defines a norm as an allocation of rights in which control over a target action is held by actors other than those who might engage in that action [Coleman, 1990, p. 243]). Here, an actor “holds a right only if it has been accorded to him by the relevant others or if it has been transferred to him by another who properly held it (and also had the right to hold it)” (Coleman, 1990, p. 58).

This elementary “theory of social action” could also explain the observed asymmetry between relative blame and praise ratings in cases of negative and positive side effects (Cova, 2016; Knobe, 2003). The reason is that the concept of a legitimate individual action does not necessarily imply any obligations of the actor to third parties but still implies a moral or legal duty not to injure those who fall within a scope of foreseeable costs of actor’s conduct if a relevant norm exists. Although some general norms may universally apply to everyday evaluations of an agent’s behavior, like avoiding foreseeable harmful consequences of planned actions to others in the short view, one can expect that different institutional domains may differ in terms of domain-specific social norms which influence people’s evaluations of very similar side effects of actions (see also Uttich & Lombrozo, 2010).

We claim that it is important to consider on theoretical level not just isolated norms related to negative side effects for some actors involved but rather institutional domains (spheres) as distinct multilayered physical, temporal, social, and symbolic sites of interaction. These domains shape norms and, more generally, life-worlds of individual and collective actors (Abrutyn, 2013, 2016). The complex of normative expectations defines not only an individual’s role during interaction but also other reciprocally related roles, for example, “doctor—patient” or, in our case, “tortfeasor—victim—arbitrator” (Parsons, 1951). Therefore, the norms regulating a specific encounter potentially leading to adverse consequences for one party cannot be applied in isolation. Blame, responsibility, or acquittal judgments of the parties involved are made within a specific institutional domain which shapes conflict behaviors according to particular subcultural scripts and underlying regulations that take into account individual or group interests. For example, a similar incidental damage suffered by a competitor might be considered as legitimate in a case of boxing fight and as blameworthy in a case of culinary contest. Hence, the microlevel processes of normatizing such encounters with respect to rules of framing, attributing categories, expressing feelings, and so on, need to be analyzed as embedded into specific institutional domains (J. H. Turner, 2001; see also Sciortino, 2005, p. 129). We are basing here on this recently refined sociological approach to institutional domains as complex social forms comprising embodied structures of roles and associated norms in comparison to more individualistic approaches which treat institutions as norms, convention, and rules intentionally constituted or situationally discovered and maintained by interacting individuals to solve emerging coordination and cooperation problems (e.g., Ellickson, 1991; D. Lewis, 1969; Opp, 2002). Taking into account specific institutional domains is essential for analyzing specific normative contexts of blaming and claiming. In other words, institutional domains may have highly specific logics of action legitimation/justification, relevant value principles, and metrics of relative importance of different people and objects potentially affected by actions’ side effects, that is, “orders of worth” (Boltanski & Thévenot, 1999; Dequech, 2008), which influence the causality and blame judgments (see also Schein & Gray, 2018 on possible effects of norm violations on harm perception). Thus, our general theoretical hypothesis states that negative side effects of intentional actions comparable in terms of foreseeability, location in causal chain, type and severity of harm caused,
and so on, may lead to very dissimilar evaluations and, therefore, different domain-specific societal reactions. Previous research on folk psychology of everyday attributions of cause and blame yielded many important findings concerning effects of such factors as blame attribution as intentionality, location of putative causes of negative action outcomes in chains of events, and so on (see, for example, Lagano & Channon, 2008). But the possible effects of differences in institutional action spheres upon judgments of cause and blame and the possible impact of such a “social” moderator variable on strength of the relations between aforementioned cognitive factors and blame attributions have remained unexplored.

We did not have specific theoretically based expectations concerning the institutional domains chosen for the study. However, we presupposed that similar negative effects for material well-being or health of the individual actor would be estimated differently for scenarios describing interactions during the purchase of medication (the pharmaceutical zone of medical institutional domain) as compared to scenarios describing interactions during the implementation of a dress code in an organizational setting of corporation (the corporate culture zone of economic institutional domain). These two institutional domains are involved into resolution of different human concerns embedded into them, that is, correspondingly, health, and subsistence (Abrutyn, 2016, pp. 213–215). There is also a distinction in culturally stereotyped and structurally supported reactions to risks and possible adverse consequences of ordinary interactions within these two domains. The organizational settings of economic organizations in many countries, including Russia, support the classification of unintended negative consequences to health as job-related accident if they occur during working hours and on a territory of organization, including corporate dining rooms and halls (Russian Federation, 2011). The health damage in such cases is routinely covered via compensatory payments from the Social Insurance Fund, which accumulates social insurance payments from employers. Such solution is not generally available for encounters occurring in the pharmaceutical zone of medical institutional domain which could lead to higher blame estimates directed personally to probable wrongdoer in the latter case to compensate for the absence of ready-made remedy.

We can also hypothesize here that any incidental “remote” damage not immediately caused to a third person by an action originator but rather by another agent does not necessarily imply blame attribution and claims for compensation. For example, this action originator could have created a new situation (e.g., established a new law or rule) and one might doubt whether this action would cause claims for compensation because blame attribution depends on an actor’s ability to control the consequences of his or her actions, which is absent in this example (Weiner, 1995). The estimated proximity of an action to the final negative outcome and the likelihood of this outcome’s occurrence in the absence of the originating action are important elements of the culpable control model of blame attribution processes (Alicke, 2000), and these theoretical assumptions are supported by empirical evidence from the previous research (Lagano & Channon, 2008). Differently stated, if another actor intervenes in a stream of events initiated by an actor-originator’s “earlier” action, the actor-originator’s blameworthiness, and causal role diminish or are even canceled out following a decrease in the degree of his or her causal control over the action’s long-term results.

**Types of social actors.** Sociological action theories quite often presuppose corporate actors, including organizations, Verbant-type associations and primary groups as sources of intentional social action. This idea of treating organizations, collectives, and primary groups as potentially blameworthy agents is not alien to folk sociology. Some empirical evidence supports this view, for example, Waytz and Young (2012). B. F. Malle (2010) even formulates the hyperagent hypothesis: “People see jointly acting groups as even more ‘agentic,’ even more driven by subjective and rational reasons, than they see individuals” (p. 116). Although the notion of macrosocial objects (such as societies or institutions) as independent entities that may act like living organisms dates back at least to the works of H. Spencer, recent empirical studies stem from the seminal work of D. Campbell (1958), who coined the concept of “entitativity” to refer to “the degree of having the nature of an entity, of having real existence” (Campbell, 1958, p. 17). There is some empirical evidence that victims’ “entitativity” influences moral judgments as well as the support for third-party collective punishment of group members (Bartels & Burnett, 2011; Pereira & van Prooijen, 2018) and that intuitive theory of groups plays an important role in shaping judgments of collective responsibility (Lickel et al., 2001; Sacchi et al., 2009). Therefore, we have some reasons to generalize the above-outlined elementary “theory of social action” to the cases of corporate actors and even decision-making institutions such as states, legislating bodies and educational institutions. However, the question of whether there are real differences in perceiving these corporate actors (e.g., whether various properties are ascribed to them) is beyond the scope of this paper. In the current research, we expected that corporate actors and, perhaps, even institutions would be judged similarly to individual actors in terms of their blameworthiness and causal role in inadvertent damage.

**Types of damage.** In many legal traditions, a person’s right to compensation for damage resulting from a loss or impairment of property, health, life, and so on, along with the scope of tort liability, is not only defined by codified rules but also decided by the judiciary. Thus, entitlement to compensation is open to case-specific legal interpretations of the “precise meaning of damage, of negligence, of the concept of causation or compensation for pure financial damages” (Schäfer,
Such interpretations may vary depending on the particular institutional domain and type of damage. Therefore, purely financial damage may be treated by the judiciary very differently from, for example, damage to a person’s health. In turn, workplace incidents leading to health damage may be considered differently from similar damage caused by manufactured products or medical malpractice. R. Lewis and A. Morris (2012) basing on their analysis of recent developments in personal injury litigation practices in the United Kingdom, demonstrated that the operation of tort law in specific domains (road traffic accidents, clinical negligence, etc.) is “much affected by the commercial interests and economic demands of the institutions that surround it” (R. Lewis & Morris, 2012, p. 592). Legal concepts and classifications related to evolving tort law are far from clear and certain. H.-B. Schäfer concludes that the legitimate place of tort law in modern society is still under discussion, with views expressed by legal scholars ranging from an expanding deterrence system dealing with old and new hazards to a restricted system that handles classical cases only and leaves complicated risks and hazards to other social institutions (Schäfer, 2000, p. 574). However, these current legal disputes are not paying enough attention to other possible sources of the general tendency toward increasing variability in the scope and popularity of tort law, namely, ordinary people’s ever-changing judgments of types of damage and lay perceptions of justifiable boundaries between those unpleasant events that should lead to blame attribution and those that should not.

Based on our current knowledge, the question of lay people’s comparative judgments about an actor’s causal role and blameworthiness in cases of different types of inadvertent damage and mischief occurring in different institutional settings, that is, large institutional action domains, has not yet been systematically studied despite the role that such judgments play in the decisions that ordinary people—as potential victims, tortfeasors, or jurymen—make in attributing damages. In our current research, we expected to discover differences in mean blame ratings of clearly different types of damage for two distinct domains we have chosen (see below).

**Current Research**

Blaming individuals, corporate actors or even institutions for breaching legal and moral norms, including those aimed at preventing unintended collateral damage to others, seems to have been a common practice in human history since ancient times (references to it can be found as early as in Hammurabi’s Code, e.g., paragraph 53) and sometimes presents a difficult problem for legal theory when, for instance, nations should be blamed without support from a supranational legislative body (Hart, 1961, pp. 208–209).

But how do ordinary blame attributions vary in different institutional domains depending on type of originating social actor? Does the purported effect of an “originator” interact with other factors (such as the type of damage inflicted or action domain)? Our study attempts to shed some light on these little-studied questions.

Previous research has demonstrated some effects of the action location in the causal chain, foreseeability, intentionality, and actor type on judgments of cause and blame and has even revealed interactions between these factors (Guglielmo & Malle, 2010; Knobe et al., 2012; Lagnado & Channon, 2008; Malle, 2010). This study is concerned with factors that have not yet been explored in depth, particularly, in their interplay. We studied how people presented with factorial vignettes depicting everyday scenarios judge an actor’s blameworthiness and causal role depending on institutional domain where everyday scenarios are taking place (the “medical” domain, the “corporate dress-code” domain), on the action-originator type (individual, organization/group or impersonal institution such as law or dress code), “remoteness” of damage from its originator (i.e., either immediate damage to an agent performing an action in a new situation created by the originator’s action or damage to a third party) and type of negative side effect (property damage or health damage). In a block full-factorial, intraindividual experiment, we also compared two institutional domains (see below in “Method” section) to test for possible domain-specific differences in both cause and blame judgments.

**Method**

**Participants**

Seventy-five students (55 female, mean age 20.3 years, \(SD = 1.6\)) in the sociology departments of the Higher School of Economics and the State University of Humanities (Moscow, Russia) participated in this experiment.

**Experimental design, materials, and procedure**

We used vignettes constructed according to a \(2 \times 2 \times 3 \times 2\) factorial design portraying everyday situations in which some well-intended action by an individual, corporate actor or institution led to negative side effects experienced by other actors, either immediately or “remotely,” that is, through the agency of another actor intentionally acting in a newly emerged situation. The general factors were as follows: (a) \(X\) — principal action originator (individual, group, and institution); (b) \(Y\) — type of unpleasant side effect (damage to property, including monetary damage, or damage to health); (c) \(Z\) — the “remoteness” of a victim (either damage to an agent himself/herself acting in a new situation or damage to a third party); and (d) \(W\) — the institutional domain of action. One institutional domain described adverse situations that emerged during the purchase of medication (the “medical” domain), and the other described situations connected to the negative side effects of implementing a dress code in an
organizational setting (the “dress code” domain). We combined these four factors in a full-factorial design to reveal not only the main effects but also the possible interactions between factors for each of two possible institutional domains. Descriptions of the relevant combinations of general factors’ levels are given in Table 1, and examples of the vignettes from different institutional domains can be found in Appendix, part B.

We constructed 12 sets with counterbalanced sequences of the resulting 24 vignettes. Each set was split into two parts—with 6 “medical” and 6 “dress code” vignettes. The first parts were quasi-randomly assigned to participants for the first experimental session (ca. 15 minutes). The second session, which was conducted on short notice, took place 2 weeks after the first session. During the second session, each participant was presented with the corresponding remaining parts of the set. In other words, each subject rated all 24 vignettes after both sessions, and due to the fact that any vignette could be presented either in the first or the second session with equal chance, we expected no effect of time on the findings.

During each session participants were asked to peruse brief descriptions of everyday situations and answer two questions concerning the causal role and blameworthiness of an actor-originator (individual, group, and institution) similar to those used by Lagnado and Channon (2008). The answers were given on 11-point discrete graphical rating scales (with $0 = \text{not at all the cause}$ to $100 = \text{completely the cause}$ and $0 = \text{not at all to blame}$ to $100 = \text{completely to blame}$), always in the same order, viz., causality judgments first. Participants were instructed to read descriptions, giving special attention to the details (“Despite the similar structures of the described situations, their details differ . . .”) and react to each situation as a whole. An example of a possible discrepancy in the ratings of cause and blame was given before the beginning of each experimental session (see Appendix A). The experiment was conducted offline, using a conventional paper-and-pencil method.

### Results

We discovered systematic significant differences between subjects’ judgments on causality and blameworthiness for two domains, “medical” and “dress code,” where the “medical” vignettes rate higher (see Figure 1A and B) as compared to analogous “dress code” vignettes depicting similar adverse outcomes. The statistical significance of the paired samples $t$-test for comparing means varies between 0.000 and 0.035 for the 12 pairs of corresponding cause judgments and between 0.000 and 0.012 for the respective pairs of blame judgments, with only two exceptions in the latter case: for the paired conditions “institution as action originator + health-harm + damage-to-a-third-party” and “individual actor as action originator + health-harm + damage-to-an-agent,” differences in the mean blame estimates for the two domains do not reach significance. Using Bonferroni correction for multiple comparisons, we receive the corrected significance level $p = .004$ which gives us more conservative estimates of seven significant differences for 12 pairs of cause judgments and nine significant differences for corresponding 12 pairs of blame judgments. Bonferroni–Holm sequential correction method (Gaetano, 2013; Holm, 1979) gives less conservative estimates: adjusted $p$-values spanning between .000 and .048 are significant at $\alpha \leq .05$ for nine pairwise comparisons between corresponding conditions for cause judgments, while for blame judgments adjusted $p$-values spanning between .000 and .036 are significant for 10 pairwise comparisons.

A repeated-measures analysis of variance (ANOVA) was conducted to estimate the main effects and interactions of four within-subject experimental factors (when the assumption of sphericity is violated, corresponding within-participants effects are always reported using the Greenhouse–Geisser adjustment).

Before reporting the main results, we should briefly describe the results of post hoc sensitivity analysis which we
conducted to find out what effect size was our study able to detect given its sample size ($N = 75$) with a power $1 - \beta = 0.95$ and $\alpha = 0.05$ (for $\rho = 0.50$). Using GPower3 for calculating of the critical population effect size (Faul et al., 2007), we obtained a value of effect size which could be discovered as small as 0.0951669.

Figure 1. Mean scores with SE for causality (A) and blameworthiness (B) for two vignette domains—buying a medication and implementing a dress code.
The main effects of institutional domain as a factor were highly pronounced for both cause and blame estimates: \(F(1, 74) = 63.758, p < .0005, \text{partial eta squared} = .463,\) where partial eta squared is a measure of effect size) for cause estimates and \(F(1,74) = 45.045, p < .0005, \text{partial eta squared} = .378\) for blame estimates.

Pairwise contrasts between both cause and blame estimates for two domains (factor W) were found to be statistically significant with mean difference between cause ratings for “medical” and “dress code” vignettes based on estimated marginal means equaling to 17.922 (SE = 2.245, \(p < .0005,\) Bonferroni-adjusted for multiple comparisons) and mean difference between corresponding blame ratings based on estimated marginal means equaling to 12.901 (SE = 1.922, \(p < .0005,\) Bonferroni-adjusted for multiple comparisons, here and below when applicable).

The statistically significant and pronounced main effect of type of damage (factor Y) was revealed for cause estimates: \(F(1, 74)=29.313, p < .0005, \text{partial eta squared} = .284,\) while for blame estimates, the main effect for the type of damage only tended to be significant: \(F(1, 74) = 3.807, p = .055, \text{partial eta squared} = .049,\) with a general tendency to slightly higher ratings for property damage.

Basing on estimated marginal means, difference between cause ratings of damage to property and damage to health equals 9.100 (SE = 1.681, \(p < .0005),\) which means that the causal role of actor-originator was generally rated higher when property (including monetary) damage was inflicted as compared to health damage.

The main effects of two other factors, \(X\) (action originator) and \(Z\) (the remoteness of a victim), did not reach the level of statistical significance for our data either for cause or for blame judgments.

In a case of cause estimates, only the first-order interaction of factor \(W\) (institutional domain) and factor \(Y\) (type of damage) was statistically significant: \(F(1, 74) = 19.464, p < .0005, \text{partial eta squared} = .208.\) As can be seen in Figure 2A, the difference between cause estimates is pronounced for the vignettes describing adverse situations that emerged during the purchase of medication (the “medical” domain) and rather negligible for the vignettes pertaining to the “dress code” domain.

For blame estimates, this interaction (domain \(\times\) type of damage) was also significant, \(F(1,74) = 10.077, p = .002,\) and partial eta squared = \(.120.\) Again, the difference between blame estimates was pronounced for the “medical” domain with higher blame ratings for property damage vignettes and turned out to be practically negligible for vignettes from the “corporate dress code” domain (Figure 2B).

Besides, the first-order interaction between factors \(W\) (domain) and \(X\) (action originator type) for blame judgments turned out to be statistically significant, \(F(1.841, 136.241) = 8.845, p < .0005, \text{partial eta squared} = .107.\) The group/collective as action originator was rated higher in blameworthiness compared to both—individual and institutional actors for the “medical” domain and only slightly lower compared to both—for the “dress code” domain. The first-order interaction between factors \(X\) and \(Y\) (type of action originator \(x\) type of damage) only tended to be significant for blame estimates, \(F(1.892, 140.044) = 3.096, p = .051, \text{partial eta squared} = .040.\) The vignettes describing the property damage were rated uniformly higher than those depicting the damage to health for both individuals and groups/collectives, while institutions as action originators received practically equivalent ratings of blameworthiness for both types of damage: estimated marginal means for blame estimates for property and health damage for institutions as actors are, respectively, 17.000 (SE = 2.089) and 17.067 (SE = 2.159).

Spearman rank correlations were computed between the cause and blame judgments. For vignettes from the “medical” domain, they vary between .251 and .516 (with corresponding “p”s between .03 and .0005) in 11 out of 12 cases (one rho = .223 only tends to be statistically significant, with \(p = .055).\) For the “dress code” vignettes, the correlation values vary between 0.411 and 0.618, with \(p < .0005\) in all cases. These data are in rather good agreement with similar data for the relation between cause and blameworthiness ratings of scenarios constructed for the comparison of chains of events (intentional, unintentional, or physical) in early and late causal positions from Lagnado and Channon (2008, pp. 761–762).

**Discussion**

Our data demonstrate the dramatic difference between two institutional domains used for construction of two sets of similar vignettes: actors-originators from “medical”-related vignettes are generally estimated to be more causally effective and blameworthy compared to the “dress code”-related actors-originators. The main effects of institutional domain factor as well as pairwise between-domain differences in mean ratings revealed in our experiment support the general theoretical hypothesis formulated above: intentional actions with negative side effects which are comparable in terms of foreseeability, location in causal chain, type, and severity of harm caused, and so on, may lead to very dissimilar evaluations when these actions are taking place “inside” different institutional domains. Previous research by M. D. Alicke (1992) already demonstrated the influence of an actor’s motives and morals on people’s judgments about the actor’s primary causal role in a car accident (see also Nadler & McDonnell, 2012). However, in these studies, the moral quality of particular action motives varies jointly with the action domains themselves (for example, a drug trafficking operation versus preparing a surprise present for a parental anniversary), so the effects of both factors cannot be disentangled. In our study, the “righteousness” of primary actions was held approximately constant; all initial actions were legitimate and morally neutral, such as issuing a decree, introducing a corporate dress code, and so on. Therefore, we
could evaluate the action domain effects independently of the moral qualities of the action motives. Other things being equal, the mean estimates of the causal role and blameworthiness of actors of any type are generally higher when the side effects of their actions lead to material or health harm connected to purchasing a prescription drug compared to analogous estimates for actors who “initiate” the similar negative effects in the corporate dress code domain. These differences in mean estimates of causality and blameworthiness are not attributable solely to outcome severity because we used comparable formulations of material and health damages inflicted for analogous experimental conditions across domains (e.g., “Anna’s state of health worsened and she was then hospitalized following a heart attack” for a “medical” vignette in a case of health harm versus “Anna sustained serious burns requiring prolonged treatment” for an analogous “dress code” vignette). It should be noted here that disentanglement of the people’s moral judgments from their perceived harm estimates remains a nontrivial task for the future research as some recent findings indicate that perceived severity of harm is a highly significant predictor of moral judgments (Gray & Schein, 2016). Although, possible influences of minor differences in corresponding wordings could also contribute to our findings. Partly, the between-domain

Figure 2. Profile plots for cause (A) and blame (B) judgments.
difference can be attributed to the distinction in culturally stereotyped and structurally supported reactions to health risks and possible adverse consequences for ordinary interactions emerging in two domains. As we noted in Introduction, the organizational contexts of work in Russia support the classification of unintended negative consequences to health as job-related accidents if they occur during working hours and on a territory of organization, including corporate dining rooms. The health damage in such cases is routinely covered via compensatory payments from the Social Insurance Fund. Such stereotyped solution is not generally available for encounters occurring in the pharmaceutical zone of medical institutional domain which could lead to the observed higher blame estimates directed personally to an action originator in the case of “medical-related” vignettes.

The observed pattern in interaction effects of institutional domain with other experimental factors used in our study also indicates the striking discrepancy between two institutional domains. Specifically, our subjects displayed a high sensitivity to difference between material and health damages in “medical-related” scenarios while their cause and blame ratings for “dress-code” scenarios from corporate life did not reveal significant differences in perception of these two types of damages. Vignettes from “medical” domain were rated higher in both causal impact and blameworthiness of an actor-originator when an agent or a third-party suffered monetary losses. Conformably, the significant main effect of the type of damage which was revealed in our study for cause estimates and will be discussed in more detail below led to higher ratings of the causal role of an actor-originator if the latter inadvertently initiated a chain of events leading to material damages, including monetary losses, for other people as compared to adverse consequences to their health. These differences may be hypothetically attributed to very general differences in the institutional contexts of medicine and corporate life as approached from the “intuitive prosecutor” position that people take when making judgments on norm violations. As P.E. Tetlock puts it, “certain contexts promote leniency or punitiveness” (Tetlock, 2002).

We can offer only a preliminary and highly hypothetical explanation of these substantive differences. If one considers some interconnections between ordinary people’s intuitive blame and cause judgments and the current state of a local (in our case, Russian) legal system that somehow epitomizes these judgments (Alicke, 2008), it can be noted that the current labor legislation tends to routinely cover health losses in the workplace via compensatory payments from the Social Insurance Fund, which accumulates social insurance deductions from employers (Russian & Federation, 1998, 2011). However, bringing a lawsuit for the repairation of minor, inadvertent moral or material damage suffered during working hours does not seem to be common in contemporary Russia. From the other side, similar material damages stemming from legislative and administrative decisions and decrees in the fields of social and health policy are generally perceived as highly contestable, that is, provoking social protests and lawsuits. For example, so-called “social benefits monetization” (i.e., the replacement of free government-funded benefits, such as medical care including the provision of free medication and transportation, with cash payments) initiated by the Russian Ministry of Health and Social Development in 2005 sparked an impressive wave of protests and almost led to the resignation of the cabinet.

Some parallels can be found between this tentative “context-based” interpretation and the results obtained by Lewis and Morris in their analysis of temporal dynamics of personal injury claims in the United Kingdom for different domains in the 2000s compared to earlier periods since 1973. Those authors have discovered that “the culture of claiming in the United Kingdom is relatively weak in the clinical, employment and public contexts” (Lewis & Morris, 2012, p. 584), whereas an impressive increase, both short- and long-term, in the number of road traffic accident claims involving personal injury compared to the dynamic of reported casualties has been observed. These authors examined the major determinants of the stronger compensation culture revealed in the latter domain, including in particular, the development of “no-win, no-fee” advertising and the claims market for road traffic accidents, which is thought to involve simpler and more routinized types of claims by the new claim management companies. The supposed parallels with our findings are rather limited, partly due to differences between Russia and the United Kingdom in their local legal contexts of claiming and partly because, as also mentioned by Lewis and Morris, referring to Felstiner et al. (1980–1981), the practices of blaming and claiming are interconnected but separate. However, it seems reasonable to expect that current claim rates in different institutional domains may influence people’s propensity to blame, and the exact mechanisms and scope of such an influence deserve further investigation.

We can offer yet another explanation of the higher blame ratings in the “medical”-related vignettes that is not necessarily incompatible with the previous explanation. In our “medical” vignettes, corporate and group-level actors are related to bodies of state power (e.g., the Minister of Health or a fraction of deputies, see Appendix B), whereas in the “dress code” vignettes, the actors are related to corporate bodies (the head of the corporation or the client department staff). This may lead to the confounding effect of differential trust levels on blame assessments. There is a relatively low level of trust in Russian governmental structures and political trust in general (see the results based on the European Social Survey studies, for example, Marien & Hoogh, 2011), so it is quite possible that if governmental actors are involved in an unfortunate event, responsibility will be primarily ascribed to those actors. Studying the link between trust and responsibility could be a promising direction for future research (Pellizzoni, 2005), but here it is enough to stress that our initial assumption of equally perceived moral neutrality of action in the two domains could be questioned.

Further research is necessary to determine the relative contributions from a situation itself, from the availability of
private insurance against specific hazards and from the perceived personal ‘harmfulness’ of a negative outcome in a specific domain. For doing so, it will be necessary to collect additional data on people’s ratings of the harmfulness and seriousness of specific negative side effects. What compensation for damage is considered fair by ordinary people and whether attributions of blame in different domains imply similar liability according to people’s judgment are other promising avenues for future research.

We also have discovered the pronounced main effect of the type of damage leading to higher ratings of scenarios describing property damages as compared to health damages for cause judgments (this effect only tended to be significant for blame judgments). The first-order interaction effect between factors of type of damage and institutional domain revealed the common pattern: the difference between ratings for property damage vignettes compared to health damage vignettes was larger for the ‘medical’ domain with higher cause and blame estimates for property damage and turned out to be practically negligible for similar vignettes from the ‘corporate dress code’ domain. Financial losses occurring as a by-product of health policy–related decrees and decisions provoked much more blame and cause attributions among our participants than unintended negative consequences on someone’s health (i.e., being compelled to pay approximately 10 times more for a substitute for a prescribed medicine versus suffering a heart attack). Hypothetically, this difference can be attributed to a greater variety of alternative causal chains with an equifinal result in the case of health damage (a heart attack can develop due to equally verisimilar attended causes such as unfavorable weather conditions, ‘natural’ rapid deterioration of coronary heart disease symptoms, and so on, not just the impossibility of timely obtaining a necessary medicine due to a new decree regulating the sale of prescription drugs). Seemingly, equally plausible causal alternatives to the chain “impossibility of buying a customary medication without prescription” \( \rightarrow \) “financial losses due to necessity to buy the much more expensive medication” are less numerous. However, our current data support this possible explanation only indirectly, especially taking into account significantly lower ratings of blame and causality for the health-harm effects in dress code vignettes, which are much more unambiguously following a potentially culpable action and approximately equaling to corresponding ratings for material losses.

We did not discover a main effect of factor \( X \), an action originator, neither for cause nor for blame judgments. These findings are in line with our general theoretical expectation that corporate actors (groups/collectives) and, perhaps, even institutions would be judged similarly to individual actors in terms of their blameworthiness and causal role in inadvertent damage.

However, the first-order interaction between factors \( W \) (institutional domain) and \( X \) (action originator type) for blame judgments turned out to be statistically significant.

The group/collective as action originator was rated significantly higher in blameworthiness compared to both—individual and institutional actors for the ‘medical’ domain (and only negligibly lower compared to both—for the “dress code” domain). This result can be considered as a partial support for the hyperagent hypothesis (Malle, 2010) which asserts that jointly acting groups/collectives could be perceived as even more “agentic” and goal-directed than individual actors. The observed between-domain difference in blame attribution to group/collective actors could be hypothetically explained by possible differences in perceived entitativity and cohesiveness of different groups (in our study, “Client Department staff” or “the fraction of deputies”) which were already shown to influence attributions of group mind and group responsibility (Waytz & Young, 2012). This assumption can be tested in future studies which may use independent measures of different actor properties (e.g., of intentionality or perceived “groupness” of collective actors in various scenarios).

Taken together, our findings provide some support to the aforementioned hypothesis of lay people’s perception of groups, corporate actors in general and, perhaps in some cases, even institutions as potentially blameworthy and causally efficient “entities.” However, they also indicate that much more should be done to specify the scope and limits of this class of phenomena.

We hypothesized that increasing ‘remoteness’ of collateral damage from an action originator would lead to diminishing people’s attributions of cause and blame. However, we have not discovered the statistically significant effects of this factor in our data. At present, there are contradictory findings on primacy and recency effects of, correspondingly, earlier or later events contributing to causation. Although there is some evidence in support of primacy effect for voluntary human actions as compared to natural physical events in opportunity chains (for a review, see Lagnado & Channon, 2008; McClure et al., 2007). Basing on our current data, we can only contemplate a possibility that actions of action originators in our vignettes were not perceived sufficiently intentional and voluntary to give rise to pronounced differences between causal and blame judgments for earlier versus later positions in chains of events leading to adverse side effects.

Thus, our findings contribute to a better understanding of lay attribution of causality and blame. We believe these results also may have practical implications because responsibility ascriptions are thought to be an important factor influencing societal responses to various adverse side effects of actions or compensation claims. Our study proves that along with psychological factors also “folk sociology”-related factors, namely institutional domain and type of actor, may influence cause and blame judgments. Therefore, taking into account the institutional domain and type of actor may be helpful in predicting the mentioned reactions to side effects of actions, compensation claims, or sanction.
requirements. However, future research is needed to clarify the link between these judgments and behavioral aspects.

Conclusions

To summarize, we found pronounced differences between ordinary people’s cause and blame estimates for two institutional domains in both the mean estimates for analogous vignettes and the patterns of main effects and interactions related to the supposed socially structured determinants of these rates, that is, type of damage and types of actors originating the unintended negative side effects. We propose some tentative alternative explanations for these findings, but future research is necessary to examine the plausibility of these explanations and to overcome the limitations of our data. The revealed differences between the two institutional action domains in effects of social factors that influence everyday attributions of cause and blame, suggest the necessity of further checks of the possible moderation effect of institutional action domain as a categorical variable on relationships between both well-known folk-psychological cognitive (intentionality, “remoteness,” etc.) and relatively underexplored social (i.e., type of actor and type of damage) factors and the resulting lay judgments of cause and blame.

Appendix

A. Instructions to Participants

Please read the following descriptions of everyday situations and answer a few questions. Despite the similar structure of the described situations, their details differ. Some elements in these descriptions will be repeated, but you must answer the questions about each situation as a whole.

After reading the description of the situation, we will ask you some questions about responsibility and causality.

Note: For example, if a 3-year-old child plays with matches and sets fire to a house, his act is the cause of the fire, but the child is not to blame for it.

When replying, please listen to your inner feelings and your own opinions regarding the cause of the damage in the situation described and the actor’s degree of responsibility.

B. Factor Level Descriptions and Example Scenarios

Table B1. Factor Level Descriptions and Example Scenarios.

| Factor levels | Type of damage | Side effect for |
|---------------|----------------|-----------------|
| Action originator (X1) | Type of damage | Institutional domain |
| Individual actor (X1) | Damage to property (Y1) | An agent, “immediate” victim (Z1) |
| Group/collective/corporate actor (X2) | Damage to health (Y2) | A third-party, “remote” victim (Z2) |
| Institution (X3) | | Medical (W1) |

Description of factor levels for the “medical” domain (W1)

(X1) Minister of Health issued a decree . . .
(Y1) had to buy the drug substitute, which was approximately 10 times more expensive.
(Z1) Anna was affected because she could not buy the medication without a prescription at a local pharmacy.

(X2) The fraction of deputies secured the adoption of a law . . .
(Y2) Z’s state of health worsened and he/she was then hospitalized following a heart attack.
(Z2) Anna’s husband was affected.

Example vignette scenario for the “medical” domain (W1)
The Minister of Health issued a decree (X1) lengthening the list of drugs sold by prescription. Medication “A” was among these drugs. Mikhail, a retiree, is seriously ill, and he has to take medication “A” regularly. His wife, Anna, could not buy the medication without a prescription at her local pharmacy. As a result, Anna had to quickly run to the doctor for a prescription and back to the pharmacy for the medicine. Because of this stress, Anna’s (Z1) health worsened, and she was then hospitalized following a heart attack (Y2).

Description of factor levels for the “dress code” domain (W2)

(X1) The head of the corporation has established strict dress code rules . . .
(Y1)’s expensive suit was hopelessly flawed.
(Z1) Anna tripped over a small metal threshold in the corporate dining room and dropped her food-laden tray. As a result, Anna was affected.

(X2) Client Department staff collectively decided to adhere to strict dress code rules . . .
(Y2) sustained serious burns requiring prolonged treatment.
(Z2) Anna tripped over a small metal threshold in the corporate dining room and dropped her food-laden tray. As a result, Ekaterina, who was sitting at the nearest table, was affected.

Example scenario for the “dress code” domain (W2)
The head of the corporation has established strict dress code rules (X1). Anna went to the office wearing a tight skirt and high heels. Anna tripped over a small metal threshold in the corporate dining room and dropped her food-laden tray. As a result, the expensive suit (Y1) worn by Anna (Z1) was ruined.
Acknowledgment

The authors would like to thank three anonymous reviewers for helpful comments.

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.

ORCID iD

Kirill A. Gavrilov https://orcid.org/0000-0002-9666-9897

References

Abrutyn, S. (2013). Revisiting institutionalism in sociology: Putting the “institution” back in institutional analysis. Routledge.

Abrutyn, S. (2016). Institutional spheres: The macrostructure and culture of social life. In S. Abrutyn (Ed.), Handbook of contemporary sociological theory (pp. 207–228). Springer.

Alicke, M. D. (1992). Culpable causation. Journal of Personality and Social Psychology, 63(3), 368–378.

Alicke, M. D. (2000). Culpable control and the psychology of blame. Psychological Bulletin, 126(4), 556–574.

Alicke, M. D. (2008). Blaming badly. Journal of Cognition and Culture, 8(1–2), 179–186.

Bartels, D., & Burnett, R. (2011). A group construal account of drop-in-the-bucket thinking in policy preference and moral judgment. Journal of Experimental Social Psychology, 47, 50–57.

Bolanski, L., & Thévenot, L. (1999). The sociology of critical capacity. European Journal of Social Theory, 2(3), 359–377.

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of person as social. Behavioral Science, 3(1), 14–25.

Coleman, J. (1990). Foundations of social theory. Harvard University Press.

Cova, F. (2016). The folk concept of intentional action: Empirical approaches. In J. Sytsma & W. Buckwalter (Eds.), A companion to experimental philosophy (pp. 121–141). John Wiley.

Cova, F., Lantian, A., & Boudesseul, J. (2016). Can the Knobe effect be explained away? Methodological controversies in the study of the relationship between intentionality and morality. Personality and Social Psychology Bulletin, 42(10), 1295–1308.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. Cognition, 108(2), 353–380.

Dequech, D. (2008). Logics of justification and logics of action. Journal of Economic Issues, XLIII(2), 527–535.

Ellickson, R. (1991). Order without law: How neighbors settle disputes. Harvard University Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39(2), 175–191.

Felstiner, W., Abel, R., & Sarat, A. (1980–1981). The emergence and transformation of disputes: Naming, blaming and claiming. Law & Society Review, 15(3–4), 631–654.

Finkelstein, C. (2005). Responsibility for unintended consequences. Ohio State Journal of Criminal Law, 2, 579–599.

Gaetano, J. (2013, May). Holm-Bonferroni sequential correction: An EXCEL calculator. https://www.researchgate.net/publication/236969037_Holm-Bonferroni_Sequential_Correction_An_EXCEL_Calculator

Gray, K., & Schein, C. (2016). No absolutism here: Harm predicts moral judgment 30x better than disgust—Commentary on Scott, Inbar & Rozin (2015). Perspectives on Psychological Science, 11(3), 325–329.

Guglielmo, S., & Malle, B. (2010). Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. Personality and Social Psychology Bulletin, 36(12), 1635–1647.

Guglielmo, S., Monroe, A., & Malle, B. (2009). At the heart of morality lies folk psychology. Inquiry, 52(5), 449–466.

Hammurabi. (Circa 1760 BC). The code of Hammurabi. http://avalon.law.yale.edu/subject_menus/hammenu.asp

Hart, H. (1961). The concept of law. Oxford University Press.

Hart, H. (1968). Punishment and responsibility: Essays in the philosophy of law. Oxford University Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics, 6(2), 65–70.

Knobe, J. (2003). Intentional action and side effects in ordinary language. Analysis, 63, 190–194.

Knobe, J., Buckwalter, W., Nichols, S., Robbins, P., Sarkissian, H., & Sommers, T. (2012). Experimental philosophy. Annual Review of Psychology, 63, 81–99.

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of Intentionality and foreseeability. Cognition, 108, 754–770.

Lewis, D. (1969). Convention: A philosophical study. Harvard University Press.

Lewis, R., & Morris, A. (2012). Tort law culture: Image and reality. Journal of Law and Society, 39(4), 562–592.

Lickel, B., Hamilton, D. L., & Sherman, S. J. (2001). Elements of a lay theory of groups: Types of groups, relational styles, and the perception of group entitativity. Personality and Social Psychology Review, 12, 140–140.

Malle, B. (2010). The social and moral cognition of group agents. Journal of Law and Policy, 19, 95–136.

Marien, S., & Hooghe, M. (2011). Does political trust matter? An empirical investigation into the relation between political trust and support for law compliance. European Journal of Political Research, 50, 267–291.

McClure, J., Hilton, D. J., & Sutton, R. M. (2007). Judgments of voluntary and physical causes in causal chains: Probabilistic and social functionalist criteria for attributions. European Journal of Social Psychology, 37(5), 879–901.

Nadler, J., & McDonnell, M.-H. (2012). Moral character, motive, and the psychology of blame. Cornell Law Review, 97, 253–304.

Opp, K.-D. (2002). When do norms emerge by human design and when by unintended consequences of human action? The
example of the no-smoking norm. *Rationality and Society*, 14(2), 131–158.

Parsons, T. (1951). *The social system*. Glencoe: Free Press.

Pellizzoni, L. (2005). Trust, responsibility and environmental policy. *European Societies*, 7(4), 567–594.

Pereira, A., & van Prooijen, J.-W. (2018). Why we sometimes punish the innocent: The role of group entitativity in collective punishment. *PLOS ONE*, 13(5), Article e0196852. https://doi.org/10.1371/journal.pone.0196852

Russian Federation. (1998, July 24). Federal law “On compulsory social insurance against accidents at work and occupational diseases.”

Russian Federation. (2011, March 10). Resolution of the Plenum of the Supreme Court of the Russian Federation of March 10, 2011, N 2 “On the application by the courts of legislation on compulsory social insurance against industrial accidents and occupational diseases.”

Sacchi, S., Castano, E., & Brauer, M. (2009). Perceiving one’s nation: Entitativity, agency and security in the international arena. *International Journal of Psychology*, 44(5), 321–332.

Schäfer, H.-B. (2000). Tort law: General. In B. a. Bouckaert (Ed.), *Encyclopedia of law and economics: Vol. I. The history and methodology of law and economics* (pp. 569–596). Edward Elgar.

Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, 22(1), 32–70.

Sciortino, G. (2005). How different can we be? Parsons’s societal community, pluralism, and the multicultural debate. In R. Fox, V. Lidz, & H. Bershady (Eds.), *After Parsons: A theory of social action for the twenty-first century* (pp. 111–136). Russell Sage Foundation.

Tetlock, P. (2002). Social functionalist frameworks for judgment and choice: Intuitive politicians, theologians, and prosecutors. *Psychological Review*, 109(3), 451–471.

Turner, J. H. (2001). Can functionalism be saved? In A. J. Trevino (Ed.), *Talcott Parsons today: His theory and legacy in contemporary sociology* (pp. 101–127). Rowan & Littlefield.

Turner, J. W. (1936). The mental element in crimes at common law. *The Cambridge Law Journal*, 6(1), 31–66.

Uttich, K., & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition*, 116, 87–100.

Waytz, A., & Young, L. (2012). The group-member mind trade-off: Attributing mind to groups versus group members. *Psychological Science*, 23(1), 77–85.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. Guilford.