A new look at the attribution of moral responsibility: The underestimated relevance of social roles

Pascale Willemsen, Albert Newen & Kai Kaspar

Published online: 19 Feb 2018.

Submit your article to this journal ⌇

View related articles ⌇

View Crossmark data ⌇

Routledge
Taylor & Francis Group

Check for updates

A new look at the attribution of moral responsibility: The underestimated relevance of social roles

Pascale Willemsen[a], Albert Newen[a] and Kai Kaspar[b]

[a]Institute for Philosophy II, Ruhr-University Bochum, Germany; [b]Department of Psychology, University of Cologne, Cologne, Germany

ABSTRACT

What are the main features that influence our attribution of moral responsibility? It is widely accepted that there are various factors which strongly influence our moral judgments, such as the agent's intentions, the consequences of the action, the causal involvement of the agent, and the agent's freedom and ability to do otherwise. In this paper, we argue that this picture is incomplete: We argue that social roles are an additional key factor that is radically underestimated in the extant literature. We will present an experiment to support this claim.

1. Introduction

Morality strongly influences our behavior, enables predictable social interaction, and thereby stabilizes a functioning society. The importance of morality comes with two core questions: what is its evolutionary source (FitzPatrick, 2016; Machery & Mallon, 2010) and what are the central features shaping our moral intuitions and moral judgments? In this paper, we will focus on the latter question by developing a new perspective concerning the relevant factors shaping our attribution of moral responsibility for an agent's actions.

The intense investigations over the previous decades have demonstrated that the attribution of moral responsibility—in the sense of attributing blame or praise to a person for an action or omission—can only be accounted for by considering various factors, such as the *agent's intentions* (Branscombe, Owen, Garstka, & Coleman, 1996; Cushman & Greene, 2011; Kant, 1785; Solan, 2002; Young, Cushman, Hauser, & Saxe, 2007), *the consequences of the action* (Bentham, 1789; Cushman & Greene, 2011; Cushman & Young, 2009; Greene, Nystrom, Engell, Darley, & Cohen, 2004; Mill, 1861; Sidgwick, 1874; Singer, 1980; Waldmann &

Wiegmann, 2010), *the causal involvement of the agent* (Alicke, 1992; Driver, 2007; Hitchcock & Knobe, 2009; Willemsen & Reuter, 2016), *and the agent's freedom and ability to do otherwise* (Nichols & Knobe, 2007; Parfit, 1984; Sinnott-Armstrong, 1984; Viney, Parker-Martin, & Dotten, 1988; Viney, Waldmann, & Bachilon, 1982).

While previous empirical studies provided good estimations of how lay people attribute moral responsibility, this picture misses an important piece. In all ethical, and the majority of psychological, approaches to moral responsibility, the agent's action is considered independently of the social context in which it is embedded.[1] However, some psychologists have argued that making moral judgments about others is an important social phenomenon that cannot be examined independently of the social context (Fiske, 1991; Gigerenzer, 2015; Malle, Guglielmo, & Monroe, 2014; Rai & Fiske, 2011). It is this insight that motivates our empirically informed argument for the relevance of social roles.

1.1. *Status quo of empirical research on the relevance of social roles*

While this is an important insight, researchers so far have failed to systematically test how exactly this social context impacts moral judgments. Some evidence in this direction is provided by Hamilton and Sanders (1981). They argue that moral judgments are sensitive to social roles, such that authorities are held to be more morally responsible than agents with lower social status. According to their view, social roles are defined by a set of norms that is specific to this social role. Thus, one and the same behavior might be judged quite differently depending on who performed it and what role-specific expectations were violated. Haidt and Baron (1996) provide evidence supporting this assumption. They found that participants who read a story about a superior performing a harmful action judged said action to be much more blameworthy than those participants who read the same story about an employee.

While some authors in psychology argue that social roles matter for moral responsibility attribution, this insight remains undetermined and has not entered the philosophical discussion. Over the last three decades, a variety of philosophically motivated experiments provided us with a better understanding of moral cognition and how moral cognition relates to normative–philosophical accounts of morality. However, the influence of social roles and hierarchies has been largely overlooked. This fact might be particularly problematic when the stories used in vignette studies feature two agents who jointly make a decision, such as in Knobe's famous CEO vignette (2003). If the two agents in such vignettes differ in their social roles, and if social roles have an effect on moral judgments, then researchers need to integrate such a confounding variable when attempting to explain experimental effects.

In a recent paper, Kaspar and colleagues (2016) show that social roles strongly affect moral judgments about a situation in which two agents jointly implement a

company activity. Their experiments demonstrate a clear asymmetry between the amount of praise and blame attributed to the boss compared to the employee, such that the employee is always attributed significantly less blame, but more praise, for consequences of the joint activity; and this effect is shown in two cultures, namely Germany and United Arabic Emirates.

1.2. Conceptual clarifications

In the following, we will use a rather rough and ready definition of the term "social roles". A social role, according to this definition, exists in a specific, institutionalized context, such as a company, a family, sportsteam, and so on. Such roles are defined by:

(i) a set of possible role-specific activities that the agent is permitted or required to perform

(ii) authority relations with respect to other social roles in the same institutional structure

(iii) a set of moral expectations of how an agent embodying the social role should act, given their role-specific activities.

According to the first condition, social roles are a network of rights and responsibilities within a social context, such as an instituitional setting. In the following, we will use two specific social roles, namely of being a manager and being a technician. The role of being a manager of a company is defined by various responsibilities that are unique to this particular role and differ from being an technician or a stakeholder. While a manager has the right and the responsibility to make far-reaching decisions on the company's future development, the technician has much more constrained responsibilities, for example, to order relevant material for repairing a machine. According to the second condition, these role-specific rights and responsibilities create hierarchical differences among different social roles. A manager's decisions are, to a very large extent, binding for other members of the company, while an technician's decision could be undone by someone higher in hierarchy. Finally, as a result of the role-specific range of activities an agent can perform, we form moral expectations as to how a manager or technician should conduct those activities. Agents with a large range of actions typically have far-reaching influence on other agents. For this reason, we expect high-status agents to take those people's interest into consideration, that is, at least to take care of any negative effects of their decision or activities for the general public.

The conceptual specification we offer here is, of course, still vague. Future research will have to further sharpen it by investigating which of the three defining features of a social role impact moral judgments and how strongly they do so. Also, additional features may be added. We will outline some ideas of how this could be done in the general discussion. However, for the purpose of the study

at hand, our specification should suffice, as we used two specific social roles that clearly differ in their respective rights and responsibilities, the hierarchical status of the two agents, and, arguably, the moral expectations we have about their respective behavior.

1.3. Aim of this paper and predictions

In this paper, we build on the work by Kaspar and colleagues, who claim that the asymmetry in moral attribution is caused by the protagonists' social roles. While this explanation accounts well for the empirical results, we still need to exclude an alternative explanation. It might be argued that the relevant factor is not the social role itself but how strongly the protagonist is causally involved in the joint activity due to social role. In Kaspar and colleagues' scenarios, it is always the boss who makes the decision to start a new activity which leads to positive or negative side effects. As a consequence, participants might have considered the boss to be much more causally responsible for the side effects. Causal responsibility is known to be an important modulator for the attribution of praise and blame (Alicke, 1992; Hitchcock & Knobe, 2009). An agent who is more involved in bringing about an outcome is typically judged more blameworthy for negative effects, all other things being equal. It is therefore reasonable to assume that the asymmetry in blame and praise that is ascribed to the boss and an employee can be *fully* explained by a perceived asymmetry in their causal involvement without any appeal to the agents' different social roles.

In order to decide between these two competing explanations, we conducted an experiment in which decision-making, and thus the most relevant aspect of causal responsibility, and social role (being the boss vs. the technician) were separated. Consequently, two alternative explanations for the asymmetry were tested.

According to the *Causal Involvement Explanation* (*CIE*), the results can be fully explained by the perceived difference in the boss's and the employee's causal involvement. Thus, reversing the decision-maker's role should result in reversed moral as well as causal judgments. More specifically, the technician will be assigned more causal responsibility as well as more blame than the boss. The *Social Role Explanation* (*SRE*) accepts causal involvement as an important and strong modulator of moral judgments. However, the SRE predicts that causal involvement is not sufficient to explain the observed effects. It rather claims that a substantial portion of the effect depends on the social role of boss and technician. Therefore, it is predicted that reversing the causal involvement should have a minor effect on moral judgments but a strong one on causal judgments. The technician will be assigned more causal responsibility when he or she makes the decision than when the boss makes it, but the boss will still be attributed much more blame than the technician.

Note that the causal involvement of an agent is not identical to his social role. While it is prototypically the case that agents in high-status social roles may start more causal chains compared to members of low-status social roles, the actual causal involvement of an agent might, however, not reflect these abilities. In our reversed scenarios, the manager keeps his social role in the company; however, in this concrete situation, he steps back from making the decision. Thus, in this particular situation, the agent's causal involvement for the company activity is reduced compared to the case in which the manager makes the decision himself.

2. Method

Overall, 209 participants (129 male) with a mean age of $M = 32.41$ years (SD = 9.80) participated in an online experiment. They were recruited on Amazon's MTurk and received monetary compensation (\$ 0.20; average time per assignment 2 min 28 s).

We applied a 2 (social role: manager vs. technician) × 2 (decision-making: normal vs. reversed) × 2 (valence of side effect: positive vs. negative) mixed-measures design, with decision-making and valence of side effect as between-participants factors and social role as a within-participants factor. Thus, we created four different versions of a vignette. The vignette that addressed a negative (in brackets the positive) side effect and that described normal decision-making (the manager makes the decision) read as follows:

> A technician of a company talks to the company's manager and presents the following idea: "The economic situation of our company is difficult. We need to check where we can reduce costs. One option might be to improve the manufacturing process. I heard about this new spare part that lasts longer and is much cheaper. However, it also leads to higher (*lower*) emission rates. But I don't care about harming (*helping*) the environment. I also see potential for economization in other divisions of the company. It's on you to decide whether we switch to the new spare part or not."

> The manager responds: "From a technical perspective, the spare part will work just as good as the more expensive one and it will certainly allow us to save a lot of money. I therefore suggest to install the new spare part. Reducing costs is all that matters to me. I don't care about harming (*helping*) the environment."

> After the manager's decision the cheap, longer lasting spare part is installed and the environment is harmed (*helped*) by higher (*lower*) emissions. The company successfully manages to save money.[2]

Depending on the specific vignette, particpants indicated how much they agree with the statements that manager and technician caused the environment to be harmed or helped, using a seven-point Likert scale from 1 (*completely disagree*) to 7 (*completely agree*). They also indicated how much blame (or praise) the manager and technician deserve for harming (or helping) the environment (1 = *no blame/ praise at all*; 7 = *very much*).

3. Results

3.1. Moral responsibility

We analyzed the attribution of blame (negative side effect condition) or praise (positive side effect) by means of a 2 (social role) × 2 (decision-making) × 2 (valence of side effect) mixed-measures ANOVA. We found main effects of social role, $F(1, 205) = 4.326$, $p = .039$, $\eta_p^2 = .021$, and valence of the side effect, $F(1, 205) = 171.374$, $p < .001$, $\eta_p^2 = .455$, but no main effect of decision-making, $F(1, 205) = 1.000$, $p = .318$, $\eta_p^2 = .005$. We also found a two-way interaction between social role and valence of the side effect, $F(1, 205) = 57.845$, $p < .001$, $\eta_p^2 = .220$, but no interaction between social role and decision-making, $F(1, 205) = 0.772$, $p = .380$, $\eta_p^2 = .004$, and no interaction between decision-making and valence of the side effect, $F(1, 205) = 0.002$, $p = .961$, $\eta_p^2 < .001$. However, the three-way interaction reached statistical significance, $F(1, 205) = 5.795$, $p = .017$, $\eta_p^2 = .027$, depicted in Figure 1.

When the technician made the critical decision eliciting a negative side effect, the difference between manager and technician decreased but was still highly significant. The manager received more blame than the technician in all conditions. Also, the amount of attributed blame to the manager was independent of the decision-making. In contrast, the technician appeared to be slightly more blameworthy when he made the critical decision, compared to the scenario in which the manager made the decision. In the positive side effect condition, the technician received more praise than the manager independently of the decision-making, but the effect was more pronounced when the manager decided.
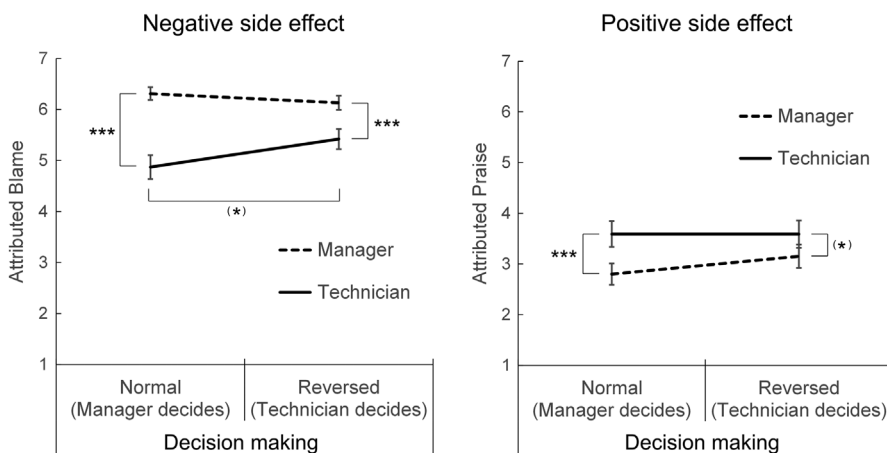
Figure 1. Attributed moral responsibility.
Notes: Vertical lines indicate the standard error of the mean. (*)$p < .10$, ***$p < .001$.
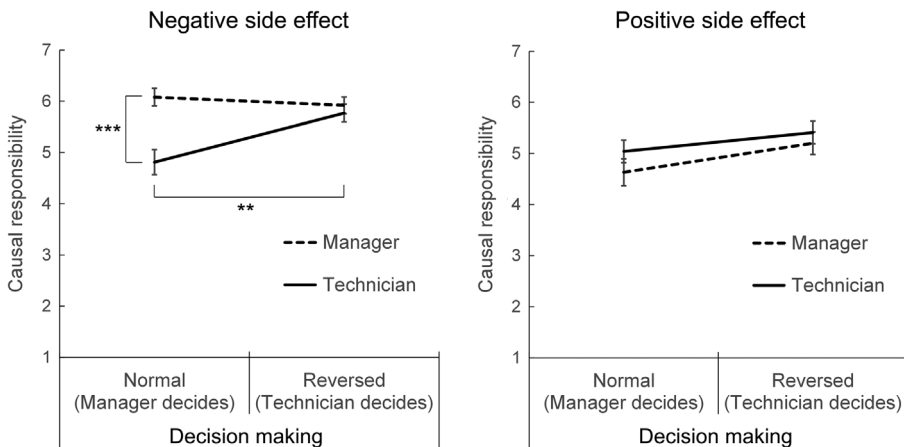
3.2. Causal responsibility

We analyzed the attributed causal responsibility by means of the same ANOVA and found main effects of social role, $F(1, 205) = 4.244$, $p = .041$, $\eta_p^2 = .020$, decision-making, $F(1, 205) = 5.226$, $p = .023$, $\eta_p^2 = .025$, and valence of the side effect, $F(1, 205) = 9.199$, $p = .003$, $\eta_p^2 = .043$. Moreover, we found two-way interactions between social role and valence of the side effect, $F(1, 205) = 20.772$, $p < .001$, $\eta_p^2 = .092$, and between social role and decision-making, $F(1, 205) = 5.253$, $p = .023$, $\eta_p^2 = .025$, but no interaction between decision-making and valance of the side effect, $F(1, 205) = 0.006$, $p = .937$, $\eta_p^2 < .001$. Most importantly, as shown in Figure 2, we again found a three-way interaction, $F(1, 205) = 8.380$, $p = .004$, $\eta_p^2 = .039$.

When the manager made the critical decision that produced a negative side effect, participants attributed more causal responsibility to the manager compared to the technician. In contrast, when the technician made the critical decision eliciting a negative side effect, no significant difference between the causal responsibility of manager and technician was found. Moreover, the causal responsibility of the manager was independent of having transferred the decision-making to the technician. In contrast, the technician was considered more causally responsible when making the critical decision, compared to the scenario in which the manager made the decision. For a positive side effect, the judgment of causal responsibility was independent of the decision-making and the social role.

3.3. Correlation between causality and moral responsibility

We finally calculated the correlation between the judgment of causal responsibility and the judgment of moral responsibility in terms of blame or praise, respectively.

Figure 2. Attributed causal responsibility.
Notes: Vertical lines indicate the standard error of the mean. **$p < .01$, ***$p < .001$.

In the case of a positive side effect, the attributed causal responsibility correlated positively with the amount of attributed blame for both the manager, $r = .611$, $p < .001$, and the technician, $r = .734$, $p < .001$. In the case of a positive side effect, causality judgment and attributed praise also correlated positively for both the manager, $r = .316$, $p < .001$, and the technician, $r = .390$, $p < .001$. However, the correlations found in the negative side effect condition were significantly larger than the correlations found in the positive side effect condition, both $Zs \geq 2.73$, $ps < .001$. Thus, attributions of causal and moral responsibility were positively related on the level of individual participants, whereas this relation was more pronounced in the case of blameworthy negative side effects.

The most important result concerning our predictions can be summarized as follows: Despite the fact that manager and technician have basically the same causal responsibility in the case of reversed decision-making, the attribution of blame remains significantly higher for the boss. This high moral responsibility of the manager can best be explained by the social role of a manager, that is, there is a socially expected high level of moral responsibility concerning activities of the company which is coupled to the social power or authority of a manager. In this respect, our results foster the Social Role Explanation (SRE).

4. Discussion

4.1. Causal involvement or social roles?

We replicated the initial findings by Kaspar and colleagues (2016). In those conditions in which the boss makes the decision to implement the new spare part, the boss is ascribed significantly more blame than the technician when the outcome is negative. In contrast, when the outcome is positive, the boss is ascribed significantly less praise than the technician. That we succeeded in gaining similar results to Kaspar and colleagues allows us to directly test the two rival explanations of this effect.

The aim of the present study was to examine two rival explantions of this asymmetry in blame attribution, namely the *SRE* and the *CIE*. The CIE claims that, all other things being equal, the agents' causal involvement can fully explain participants' blame and praise judgments. The SRE accepts that causal involvement is a relevant factor for attributing blame and praise. However, in addition to the agents' causal involvement, their respective social roles are considered a further, irreducibale factor that influences people's moral judgments. Only by adding social roles to the picture can we explain differences in moral judgments when two agents interact and jointly make a decision.

Let us start with the morally negative side effect. By reversing the decision-making, we increase the causal responsibility of the technician and, so one should think, reduce the manager's causal responsibility. If our manipulation of the causal involvement worked, we should see that participants give higher causal responsibility to the technician in the reversed case.[3] Our data show exactly this significant

increase for the employee, whereas the manager remains at the same level of causal responsibility. As a result, in the reversed decision-making condition, manger and technician are at the same level of causal responsibility. These data show that the intended modulation of causal responsibility for the technician was successful and allows to distinguish between the competing predictions.

Given the differences in causal responsibility attribution, how well can CIE and SRE account for the observed effects in people's moral judgments? The CIE claims that moral responsibility should perfectly coincide with causal responsibility and that social role is irrelevant, all other things being equal. The CIE, thus, predicts that in the reversed decision-making condition, the blame for the technician should significantly increase due to the increase in the technician's causal involvement. Since we know that causal judgments for the manager did not change significantly, we should expect no difference in the attribution of blame to the manager. Most importantly, since the technician and the manager have the same level of causal involvement in the reversed decision-making case, they should also reach the same level of blame in this condition. In contrast, SRE denies a perfect coincidence of causal and moral responsibility judgments. It predicts that in the reversed decision-making condition, although the level of blame can be slightly increased for the technician (due to the significant increase of causal responsibility), it should not reach the same level as that of the manager. The social role of the technician should partly block blame attribution such that the technician should remain significantly lower in the blame attribution compared to the manager. And this is exactly what our data show.

Let us now have a closer look for the second main condition, that is, the attribution of praise for a positive side effect. In the normal decision-making condition, both manager and technician are ascribed the same amount of causal responsibility. And yet, the technician is considered significantly more praiseworthy compared to the manager. These tendencies remain the same when we reverse the decision-making: With respect to the causal judgments, we see that there is still no difference between the two agents. With respect to the moral judgments, by trend, the technician is also ascribed more praise compared to his boss.

The results for the positive side effect raise a series of questions. On the one hand, it is not clear why manager and technician are ascribed the same amount of causal responsibility in the normal decision-making condition. The story is, with the exception of the words "harming" and "helping", completely identical to the negative side effect story in which people saw a large difference in the agents' causal involvement. In addition, the manipulation of reversing the decision-making did not affect people's causal judgments in the positive side effect condition. So far, we lack an explanation of these effects. However, what is crucial for testing the plausibility of CIE and SRE is that causal judgments do not seem to explain why the technician is ascribed more praise compared to the boss. If, as the CIE claims, causal involvement is all that matters, manager and technician should be ascribed the same amount of praise. The SRE does offer an explanation: The differences in

the moral evaluation can be explained by manager and technician having different social roles. One relevant aspect of social roles that might have played a key role here are the moral expectations that we have toward members of the two roles. Since a manager has a far wider range of rights and responsibilities, we expect him/her to conduct his/her activities in a way that he/she cares about benefits for the community (if not too costly for the company). Thus, a positive side effect that can be reached at low costs is something that we assume a manager should strive for. Benefitting the community or the environment is part of our moral expectations. However, we do not hold the same expectations toward a technician. Knowing that a technician is rather unlikely to ever be able to have a direct impact on society or the environment (at a larger scale), the expectation to care about society and environment is not definitive of the role of being a technician. As a consequence of these role-specific expectations, the manager just did as he/she was expected to do. The technician, however, excelled by doing something beyond our expectations.

The empirical results we presented in this paper lend support to the view that social roles are an additional factor influencing moral responsibility attribution. The CIE is more parsimonious, yet it fails to explain our empirical results. The following two main effects are only accounted for if, in addition to the actual causal involvement, we accept the relevance of social role. First, in the negative side effect condition, we saw that reversing the decision-making leads to a convergence of causal judgments for boss and employee. However, there remains still a significantly higher level of blame for the manager, compared to the technician. Second, in the positive side effect condition and the normal decision-making scenario, the technician is attributed significantly more praise compared to the manager, although the technician has been judged equally causally reponsible. The CIE cannot explain why these effects occur.

We consider our results an important step toward a more complete picture of moral responsibility attribution. On the one hand, they add an additional factor to extant models of moral responsibility attribution. In addition, the results further question the predominant influence of monofactorial and individualistic ethical theories. In the following, we would like to emphasize the limitations of our experimental design and the conclusions that should be drawn from them. We will outline three open questions that might direct future empirical research.

4.2. Limitations of the experimental design and outlook

First, as we have outlined in the introduction, the concept of social role as we have used it throughout this paper is a rather rough-and-ready one. Future research will have to engage more closely with the following two questions: First, we need to pin down the defining features of social roles more clearly. The list of features we provided might be a good first approximation but is unlikely to be exhaustive. A better understanding of what social roles are will enable a more thorough investigation of

how social roles affect moral judgments. Thus, second, we need more systematic research on the impact of social roles. For instance, our experiment cannot provide any evidence as to which of the three defining features actually caused the effects we found. The results might hinge entirely on the agents' relative powers to act within their instituitional setting (feature 1), the relative hierarchical differences between the agents (feature 2), or on the moral expectations we have about the agents' behavior (feature 3)—and it might equally be that all of these features are relevant to different degrees.

Future research will have to examine cases in which these features come apart. For instance, if only the agents' powers to act, that is their rights and responsibilities within the institution, are relevant, we should find no differences in moral responsibility attribution when two agents of the same social role interact, say two managers or two technicians. However, when we compare two scenarios in which agents of the same social role interact, but we manipulate whether these agents are managers or technicians, we should find a difference. Such an experiment would be a manipulation of our experimental stimuli, using a between-participant design. Alternatively, one might believe that moral expectations play the predominant role. To test this hypothesis empirically, one might want to test social roles with which we connect very different moral expectations. Examples are people working in health care, such as doctors, nurses, or midwives, where other people's well-being is a priority, in jobs that provide moral guidance to others, such as pastors, priests, nuns, rabbis, and imams, and in jobs related to environmental issues, such as environmental activists who deeply care about environmental issues. Finally, there is a way to test the relevance of the hierarchical differences among the interacting agents. We know from a variety of studies that hierarchies differ quite strongly across cultures. While Western countries are usually characterized by rather flat hierarchies, other countries are vertically organized. If hierarchical differences produce the effect, we should find that the more strongly a country interprets social hierarchies, the stronger the observed social role effect will be.

Second, it might be argued that the experimental manipulation we used only partially worked. When we manipulated the decision-making, we saw a significant increase in the technician's causal relevance. This result is fully in line with our prediction. However, the manager remains at the same level of causal responsibility. This result deserves further explanation. Why is it that our manipulation affected the technician's but not the manager's causal responsibility? It seems that participants did understand that it was now the technician who made the decision, and yet this did not reduce the manager's causal responsibility.

We know from a series of empirical studies that people heavily rely on counterfactual reasoning when they make causal judgments (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2014; Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015; Stephan, Willemsen, & Gerstenberg, 2017). According to counterfactual theories of causation, an event E caused an outcome O if it is true that if E had not happened, O would not have happened. And the same rationale applies to

omissions as well. In the reversed decision-making condition, people might have given high causal responsibility to the manager because they believed that the manager could have intervened and stopped the technician's decision from being causally effective. Thus, the outcome was not only the result of the technician's decision, but also of the manager's not intervening. In the normal decision-making condition, this counterfactual reasoning does not equally well apply to the technician, though. As we have outlined in the introduction, social roles are partly defined by what the agents can and cannot do. Typically, intervention rights are part of what it means to be a manager, whereas it is not part of being a technician. Thus, participants are likely not to have reasoned that if the technician had intervened, the manager's decision would not have been causally effective. If this explanation is actually valid, we understand why our manipulation affected the technician's causal responsibility more strongly compared to the manager. It further points in an interesting direction for future research: Social roles and their defining rights and responsibilities seem to not only affect moral judgments. They also seem to alter the counterfactuals we consider relevant for causally evaluating the scenario. There is already evidence that moral considerations and statistical information alter the relevance of counterfactuals (Phillips, Luguri, & Knobe, 2015). However, nobody has ever demonstrated that this effect is also produced by social roles.

Third, we believe that additional work needs to be done on the differences between the praise and the blame conditions. As our results demonstrate, participants evaluated both agents' praiseworthiness quite differently from their blameworthiness. While the manager received more blame compared to the technician across conditions, he was always ascribed less praise than the technician. In both the philosophical and the psychological literature on moral judgments, researchers have usually assumed that "moral responsibily" or "being morally responsible" are just umbrella terms for being either praise- or blameworthy. However, it seems that moral responsibility judgments look quite different depending on whether they express praise or blame. Thus, a more nuanced conceptualization is needed.

While we admit that more research needs to be conducted to better understand what social roles are, how they are constituted, and why they affect moral evaluation, this article clearly indicates that social roles and hierarchies need to be included into systematic research in moral cognition. Neglecting this factor can only draw a radically incomplete picture of the practice of moral responsibility attribution.

Notes

1. We do not intend to imply that the empirical evidence we will provide in this paper can immediately inform the normative debate in ethics concerning how we *should* ascribe moral responsibility or what it means to be morally responsible. It is an independent and quite controversial discussion whether this is ever possible. For the purpose of this paper, we will remain silent on this issue and aim to contribute to the work on descriptive ethics and moral psychology, rather than normative ethics.

2. Other relevant adaptions of the vignette are: In the Reversed condition, the first sentence reads: "*A manager of a company talks to a technician of the company and presents the following idea.*" Consequently, in the second paragraph "*The technician responds*," and the final paragraph starts with "*After the technician's decision.*"

3. In fact, one should also expect the manager to be attributed much less causal responsibility in the reversed decision-making condition. This effect, however, did not occur. Causal ratings for the manager did not differ as strongly as one would expect. We will discuss potential explanations of this effect below. For now, the manipulation did work sufficiently well, as we do see changes in the ascribed causal responsibility to at least one agent. This allows us to test the CIE, according to which any change in moral responsibility should be explainable by a change in causal responsibility.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

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

Bentham, J. (1789). *An introduction to the principles of morals and legislation*. Garden City: Doubleday.

Branscombe, N. R., Owen, S., Garstka, T. A., & Coleman, J. (1996). Rape and accident counterfactuals: Who might have done otherwise and would it have changed the outcome? *Journal of Applied Social Psychology, 26*(12), 1042–1067.

Cushman, F., & Greene, J. D. (2011). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience, 7*(3), 269–279.

Cushman, F., & Young, L. (2009). The psychology of dilemmas and the philosophy of morality. *Ethical Theory and Moral Practice, 12*(1), 9–24.

Driver, J. (2007). Attribution of causation and moral responsibility. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (vol. 2, pp. 423–439). Cambridge, MA: MIT Press.

Fiske, A. P. (1991). *Structures of social life: The four elementary forms of human relations: Communal sharing, authority ranking, equality matching, market pricing*. New York, NY: Free Press.

FitzPatrick, W. (2016). Morality and evolutionary biology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. New York, NY: Free Press. Retrieved from https://plato.stanford.edu/entries/morality-biology

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2014). From counterfactual simulation to causal judgment. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 523–528). Austin, TX: Cognitive Science Society.

Gigerenzer, G. (2015). *Simply rational: Decision making in the real world*. New York, NY: Oxford University Press.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron, 44*(2), 389–400.

Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology, 26*(2), 201–218.

Hamilton, V. L., & Sanders, J. (1981). The effect of roles and deeds on responsibility judgments: The normative structure of wrongdoing. *Social Psychology Quarterly, 44*(3), 237.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *The Journal of Philosophy, 106*(11), 587–612.

Kant, I. (1785). *Grundlegung zur Metaphysik der Sitten* (1978 ed.). Berlin: Akademie-Ausgabe, de Gruyter.

Kaspar, K., Newen, A., Dratsch, T., de Bruin, L., Al-Issa, A., & Bente, G. (2016). Whom to blame and whom to praise: Two cross-cultural studies on the appraisal of positive and negative side effects of company activities. *International Journal of Cross Cultural Management, 16*(3), 341–365.

Knobe, J. (2003, July 1). Intentional action and side effects in ordinary language. *Analysis, 63*(279), 190–194. doi:10.1111/1467-8284.00419.

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition, 137*, 196–209.

Phillips, J., Luguri, J. B., & Knobe, J. (2015). Unifying morality's influence on non-moral judgments: The relevance of alternative possibilities. *Cognition, 145*, 30–42.

Machery, E., & Mallon, R. (2010). Evolution of morality. In J. Doris (Ed.), *The moral psychology handbook* (pp. 3–46). Oxford: Oxford University Press.

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry, 25*(2), 147–186.

Mill, J. S. (1861). *Utilitarianism*. New York, NY: Oxford University Press.

Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Nous, 41*(4), 663–685.

Parfit, D. (1984). *Reasons and persons*. New York, NY: Oxford University Press.

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*(1), 57–75.

Sidgwick, H. (1874). *The methods of ethics*. London: Macmillan.

Singer, P. (1980). *Practical ethics*. Cambridge: Cambridge University Press.

Sinnott-Armstrong, W. (1984). 'Ought' conversationally implies 'Can.' *Philosophical Review, 93*(2), 249–261.

Solan, L. M. (2002). Cognitive foundations of the impulse to blame. *Brooklyn Law Review, 68*, 1003.

Stephan, S., Willemsen, P., & Gerstenberg, T. (2017). Marbles in inaction: Counterfactual simulation and causation by omissions. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39nd annual conference of the Cognitive Science Society* (pp. 1132–1137). Austin, TX: Cognitive Science Society.

Viney, W., Waldmann, D., & Bachilon, J. (1982, November 1). Attitudes toward punishment in relation to beliefs in free will and determinism. *Human Relations, 35*(11), 939–949. doi:10.1177/001872678203501101.

Viney, W., Parker-Martin, P., & Dotten, S. D. H. (1988). Beliefs in free will and determinism and lack of relation to punishment rationale and magnitude. *The Journal of General Psychology, 115*(1), 15–23.

Waldmann, M. R., & Wiegmann, A. (2010). A double causal contrast theory of moral intuitions in trolley dilemmas. In *Proceedings of the 32nd annual conference of the Cognitive Science Society* (pp. 2589–2594). Austin, TX: Cognitive Science Society.

Willemsen, P., & Reuter, K. (2016). Is there really an omission effect? *Philosophical Psychology, 29*(8), 1142–1159.

Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007, May 15). The neural basis of the interaction between theory of mind and moral judgment *Proceedings of the National Academy of Sciences, 104*(20), 8235–8240. doi:10.1073/pnas.0701408104.