

Variations in judgments of intentional action and moral evaluation across eight cultures

Erin Robbins¹

Jason Shepard²

Philippe Rochat³

¹School of Psychology & Neuroscience, University of St Andrews, UK

²Agnes Scott College, Atlanta, GA, USA

³Emory University, Atlanta, GA, USA

*Corresponding Author:

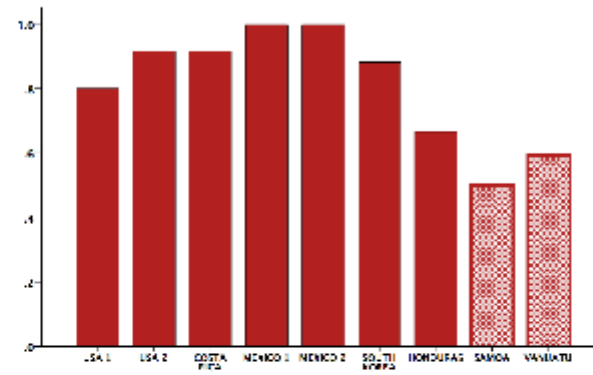
Erin Robbins
School of Psychology & Neuroscience
University of St Andrews
St Marys Quad
St Andrews, Scotland KY16 9JP
er70@st-andrews.ac.uk

In the side-effect effect, individuals tend to judge bad side effects as more intentional than good side effects (left panels), and harmful outcomes more worthy of blame than helpful outcomes worthy of praise (right panels). Across highly contrasted populations we found limited generalizability of the original side-effect effect, and even a trend toward inversion of the effect in two rural, traditional cultures (Samoa and Vanuatu) where participants were more likely to judge the helpful side effect as intentional and praiseworthy (shaded bars).

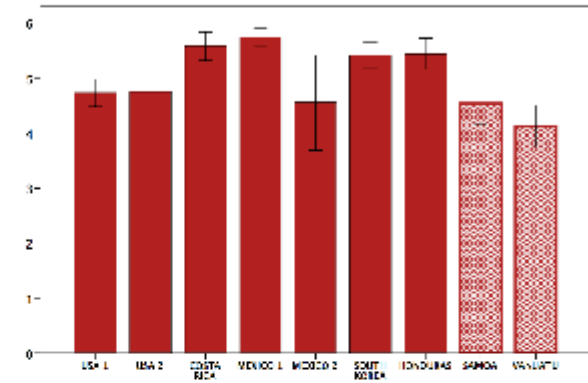


HARM THE ENVIRONMENT

% of participants who judge the unintentional harmful consequence to the environment as INTENTIONAL



Mean rating of BLAMEWORTHINESS



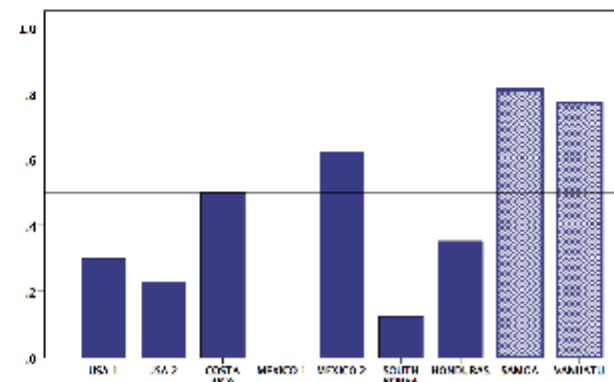
The leader of a group decides to adopt an agricultural practice that will increase community profits but *unintentionally...*

OR

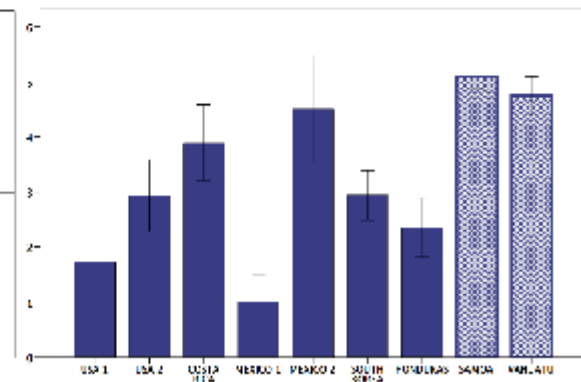


HELP THE ENVIRONMENT

% of participants who judge the unintentional helpful consequence to the environment as INTENTIONAL



Mean rating of PRAISEWORTHINESS



Abstract

Individuals tend to judge bad side effects as more intentional than good side effects (the Knobe or side-effect effect). Here, we assessed how widespread these findings are by testing eleven adult cohorts of eight highly contrasted cultures on their attributions of intentional action as well as ratings of blame and praise. We found limited generalizability of the original side-effect effect, and even a reversal of the effect in two rural, traditional cultures (Samoa and Vanuatu) where participants were more likely to judge the good side effect as intentional. Three follow-up experiments indicate that this reversal of the side-effect effect is not due to semantics and may be linked to the perception of the status of the protagonist. These results highlight the importance of factoring cultural context in our understanding of moral cognition.

Keywords: moral cognition; moral evaluation; intentional action; side-effect effect; cross-cultural psychology

1. Introduction

The ability to distinguish between intentional and non-intentional action is an essential component of social cognition (Malle, Moses, & Baldwin, 2001). In general, intentional harms are judged more harshly than unintentional harms (Cushman, 2008; Cushman, Sheketoff, Wharton, & Carey, 2013; Young, Cushman, Hauser, & Saxe, 2007). Intention-based moral evaluations and third party preferences are early developmental facts, observable in babies younger than 12 months (Hamlin, 2013; Hamlin, Wynn & Bloom, 2007; Nobes, Panagiotaki, & Pawson, 2009). Some researchers have argued that the relation between attributions of intentional action and moral evaluations, either positive or negative, is an innate principle of our moral psychology, part of a “universal moral grammar” (Mikhail, 2007). As a case in point, in U.S. criminal law and the codes of most other cultures, intentional harms tend to be judged more severely than non-intentional harms (Fletcher, 1998; Green, 2000). In the U.S., manslaughter is associated with lesser penalties (10-16 months in prison), whereas the federal sentence for murder ranges from 19.5 years in prison to a mandatory life sentence (Federal Sentencing Guidelines Manual §2A1.1-§2A1.4). The punishment is different, even though the absolute outcome of the crime is the same. Likewise, the severity of our moral judgments depends in general on our ascription of relative intention behind the offense.

Recent research in psychology and philosophy draws a complex picture of the relation between attribution of intentional action and moral evaluation. For example, numerous findings report that people are much more likely to judge that bad outcomes are brought about intentionally compared to good outcomes, the so-called side-effect effect or Knobe effect (original research by Knobe, 2003a, 2003b, 2005, 2006, 2010; Knobe & Mendlow, 2004; see additional studies by Cova & Naar, 2012; Cushman & Mele, 2008; Ditto, Pizarro, &

Tannenbaum, 2009; Lanteri, 2012; Nadelhoffer, 2004a, 2004b, 2005, 2006; Pellizzoni, Girotto, & Surian, 2010; Sousa & Holbrook, 2010; Wright & Bengson, 2009). The side-effect effect has been claimed to reflect deep and fundamental facts about human cognition. However, these claims often occur in the absence of considerations of culture and context. Our approach here is cross-cultural. The general rationale guiding our study is that if these effects are truly intrinsic and pervasive facts about our moral psychology, they are likely to be universal and should hold outside the predominantly W.E.I.R.D. (White Educated Industrialized, Rich and Democratic) populations tested to date (Henrich, Heine, & Norenzayan, 2010). The question is whether these phenomena might hold across a wide range of cultures, as implied by many moral theorists studying the side-effect effect and allied phenomena.

In the original Knobe (2003a) study, participants were presented with one of two scenarios. Both scenarios involved a decision made by the chairman of a board of a company to increase company profits. The only thing that differed between the two scenarios was the goodness or badness of a foreseen side effect of the chairman's decision, specifically whether the environment was helped or harmed as a result of the decision. After reading the scenario, participants were asked whether the chairman intentionally harmed (or helped) the environment and whether the chairman deserved blame (or praise) for harming (or helping) the environment. Knobe's results were striking: Eighty-two percent of participants said that the chairman intentionally harmed the environment, while only 23 percent said the chairman intentionally helped the environment. In response to the blame/praise question, participants strongly agreed that the chairman deserved blame for harming the environment, but that the chairman did not deserve praise for helping the environment. The author also found a strong, positive correlation between attributions of intentional action and judgments of blame/praise.

The tendency to attribute intentions to negative but not positive outcomes (the side-effect effect) has been observed across a wide variety of methodologies. This asymmetry in intentional action attributions has been replicated with other scenarios (Knobe, 2003b; Knobe & Mendlow, 2004; Mallon, 2008; Nadelhoffer 2004a, 2006; Shepard & Wolff, 2013; Uttich & Lombrozo, 2010; Wright & Bengson, 2008), with children as young as four years (Leslie, Knobe, & Cohen, 2006), with participants who suffer from deficits in emotional processing due to lesions in the ventromedial prefrontal cortex (Young, Cushman, Adolphs, & Hauser, 2006), and, for at least some versions of the scenarios, with adults with high functioning autism or Asperger's (Zalla & Leboyer, 2011; Zalla, Machery, & Leboyer, 2010). The asymmetry has also been reported with word changes in the original script introducing varying concepts such as *intention* and *intending* (McCann, 2005), *deciding*, *being in favor of*, *advocating for* (Pettit & Knobe, 2009), *knowledge* (Beebe & Buckwalter, 2010), *belief* (Beebe, 2013; Tannenbaum, Ditto, & Pizarro, 2007), *awareness* (Tannenbaum, et al., 2007) *remembering* (Alfano, et al., 2012), and *desire* (Guglielmo & Malle, 2010; Tannenbaum, et al., 2007).

In general, these findings have been interpreted to suggest that our concept of intentional action, along with other theory of mind concepts, is fundamentally moral or morally driven (Knobe, 2005) and that this role is deep and pervasive (Knobe, 2005, 2006, 2010). Further support for such moral underpinnings comes from the fact that there is a pervasive asymmetry in the ratings of blame and praise (Knobe, 2003a).

In all, the question is whether such asymmetries do actually tell us anything fundamental about the way we think about moral evaluations, intentional action, and their relation. In other words, we ask whether the side-effect effect is truly universal or might vary across cultures. To our knowledge, a limited number of cross-cultural comparisons have yielded contrary findings.

Knobe & Burra (2006) replicated the side-effect effect in a sample of 61 US college Hindi-speaking students in South Asian clubs at Princeton University and Yale University using a Hindi translation of the original script, but a more recent comparison of US participants and Indian participants tested in either Hindi or English suggests a reversed side-effect effect, whereby Indian participants judged helpful acts as more intentional than harmful ones (Clark et al., 2017). The explicit claims, and often implicit assumptions, of the side-effect effect as being intrinsic to human cognition require further scrutiny. To further address the question of how universal these effects are, we assessed them across eleven cohorts of adults amongst eight highly contrasted cultures (Study 1). In a series of follow-up experiments, we demonstrate that inversions of the side-effect effect may be linked to perceptions of status (Studies 2-4).

2. General method

A convenience sample of 464 participants (253 female) between the ages of 14-90 ($M=31.39$, $SD=13.40$ years) completed the study. Table 1 summarizes the demographic information for each of the eleven cohorts and their corresponding eight different cultures. Participants were recruited and tested individually in public spaces (e.g., public park, company break room) by experimenters who were residents of the area and fluent in the local language, or who were non-natives assisted by local informants fluent in both English and the local language. In all cultures, the vignettes and follow-up questions (described in detail below and reproduced in the Supplemental Materials) were translated and back-translated by independent research assistants who were native speakers of the local language and also fluent in English. The translated vignettes were then presented to participants in written format and read aloud by the experimenter or her assistant. At test, participants indicated their responses to the experimenter,

who recorded them on a coding sheet. Participants who could not answer prompts probing for story comprehension were excluded from analysis, yielding an attrition rate of 2% collapsed across all cultures.

Table 1: Demographic highlights of the eleven cohorts across eight cultures.

Country	Region	Setting	Environment	SES status	Population Highlights	Design
USA 1 (N= 60)	New York City	Public Space	Urban	Middle/High	Mixed ethnicity and income; Age=18-78 years, $M=38.91$, $SD=14.28$ years, 30 females	Between Subjects (Harm Condition: N = 30, Help Condition: N = 30)
USA 2 (N=25)	Metro-Atlanta	University	Urban	Middle/High	Predominantly Caucasian undergraduate students; Age=19-25 years, $M=20.71$, $SD=1.52$ years, 13 females	Between Subjects (Harm Condition: N = 12, Help Condition: N = 13)
USA 3 (N=56)	Metro-Atlanta	University	Urban	Middle/High	Predominantly Caucasian undergraduate students; Age=18-50 years, $M=20.20$, $SD=5.21$, 30 females	Within Subjects (all participants completed both help and harm conditions)
USA 4 (N=60)	Metro-Atlanta	University	Urban	Middle/High	Predominantly Caucasian undergraduate students; Age = 18-65 years, $M=24.77$, $SD=10.28$, 35 females	Within subjects (all participants completed both help and harm conditions)
SAMOA (N = 73)	Faga, Savai'i, Polynesia	Village	Rural, chief system, traditional, collectivistic and highly hierarchical (<i>matai</i>) system.	Very Low	Adults of a subsistence living village with a population of approximately 500; Age=14-65 years, $M=30.48$, $SD=13.46$ years, 38 females	Within-Subjects (all participants completed both help and harm conditions)
VANUATU (N = 67)	Motalava, Banks Island, Melanesia	Village	Rural, chief system, traditional, collectivistic and egalitarian	Very Low	Adults of a highly insular and isolated subsistence living village with a population of approximately 1000; Age=17-90	Within-Subjects (all participants completed both help and harm conditions)

					years, $M=32.56$, $SD=15.43$ years, 42 females	
COSTA RICA (N=21)	Talamanca	Village	Rural, chief system, traditional, matrilineal, collectivistic	Very Low	Adults of a traditional BriBri village depending heavily on external aide with a population of approximately 300; Age=15-70 years, $M=40.00$, $SD=15.28$, 10 females	Between-Subjects (Harm Condition: N = 9, Help Condition: N = 12)
MEXICO 1 (N=16)	Chiapas	Custapec	Rural	Low	Chiapas-born Mexican employees working and living on an isolated coffee plantation; Age=15-55 years, $M=36.31$, $SD=12.61$, 6 females	Within-subjects (all participants completed both help and harm conditions)
MEXICO 2 (N=17)	Chiapas	Custapec	Rural	Low/ Very Low	Guatemalan migrant workers seasonally employed and temporarily living on an isolated coffee plantation; Age=15-55 years, $M=25.00$, $SD=11.60$, 7 females	Within-subjects (all participants completed both help and harm conditions))
SOUTH KOREA (N= 34)	Seoul	Public Space	Urban	Middle/High	Adults of mixed income; Age=21-61 years, $M=36.21$, $SD=10.79$ years, 21 females	Between-subjects (Harm condition: N = 18, Help Condition: N = 16)
HONDURAS (N=35)	San Pedro Sula	Workplace	Urban	Middle/High	Adults of a Korean-owned company working in Honduras; Age=21-57 years; $M=34.43$, $SD=8.46$, 21 females	Between-subjects (Harm-condition: N = 18, Help Condition: N = 17)

Note: Experiment 1 included all populations but the USA 3 and USA 4 cohorts; only a subset of participants in Samoa (N=27) and Vanuatu (N=34) additionally completed Experiment 2. Experiment 4 included only participants from the USA 4 cohort.

In two conditions (harm versus help), we assessed whether judgments of intentional action, as well as ratings of either blame or praise, depended on the framing of a story in which a protagonist makes a decision about an agricultural policy that will increase community profits, but with consequences for the local environment and crops. This protagonist (either a CEO or village High Chief depending on culture) represents an individual recognized as having decision-making power and social ascendance. They represent a supreme authority in economic decisions. In this sense, the responsibilities of CEO and Chief overlap. In the vignettes, a subordinate approaches the protagonist with a suggestion about how to increase profits. In the “harm” condition, this suggestion will increase profits but also damage the environment. In the “help” condition, this suggestion will increase profits and also improve the environment. In both conditions, the protagonist responds by saying that his only concern is maximizing profit, and that he does not care at all about the effect on the environment. The suggestion is adopted, and the vignette ends with the environment being either harmed or helped. Crucially, the vignettes differ only in the effect (help or harm) that the suggestion will have on the environment (see Supplemental Materials). Note that the well-being of the local environment was a moral concern for all cultures, and particularly so for our subsistence-based populations (see O’Meara, 1990, Shore, 1982, and Vienne, 1984 regarding the South Pacific but also Cusack & Dixon, 2006, Dahlquist et al., 2007, and Polidoro et al., 2008 regarding environmental practice in Central America).

Following the story, participants answered two questions. The first question asked participants to determine how much praise (help condition) or blame (harm condition) the protagonist deserves on a seven-point Likert scale, with zero indicating none and six indicating a

lot. Across culture and conditions, participants tended to make full use of the scale. The second question asked participants whether they thought the leader *intentionally* helped or harmed the environment (coded as yes or no). Although other versions of the paradigm have used continuous measures for this judgment of intentional action, we elected to use a dichotomous outcome because intentional action is typically construed as binary (i.e., either someone acted intentionally or they did not) and in order to keep the question simple by avoiding an agree-disagree scale.

3. Experiment 1

3.1 Method

In Experiment 1 (N= 348 participants), the decisions described in the vignette were enacted by a protagonist described as either the CEO of an organization (USA 1, USA 2, Mexico 1, Mexico 2, South Korea, and Honduras) or alternately as the High Chief of a village (Samoa, Vanuatu, and Costa Rica) to reflect local leadership conventions (see Supplemental Materials). Half of the participants (N=175 from five populations: USA 1, USA 2, Costa Rica, Honduras, South Korea) received only one condition (i.e., either help or harm). The other half of participants (N=173 from four populations: Mexico 1, Mexico 2, Samoa, and Vanuatu) were tested successively in both conditions (help and harm) in a counterbalanced order. This within-subject manipulation opportunistically allowed us to assess whether potential asymmetries between judgments of intentional action and praise/blame ratings noted at a group level might also hold within individuals.

3.2 Results

In an initial analysis, and in order to compare the nine populations directly, we considered only responses to the first vignette read by the participant (harm or help). We first compared the proportion of participants who judged the protagonist's action as intentional. Results indicate a significant joint influence of culture and condition (help or harm) on judgments of intentionality ($\chi^2 = 14.37$, $df = 1$, $p < .001$, Cramer's $V = .167$). In follow-up tests, the majority of cultures demonstrated either a significant (USA 1, USA 2, Mexico 1, South Korea; all $p < .01$ based on Fisher's exact tests) or marginal tendency (Costa Rica, Mexico 2, Honduras) to ascribe intentionality to the protagonist in the harm but not the help condition (Figure 1). This asymmetry was inverted significantly in Samoa ($p < .01$) and to a lesser extent in Vanuatu ($p = .096$) where participants tended to judge the protagonist's actions as intentional in the help but not the harm condition. These results indicate that the original Knobe report of an asymmetry in judgments of intentionality between harm versus help conditions does not hold in our South Pacific samples where attributions of intentional action are more common in the help condition.

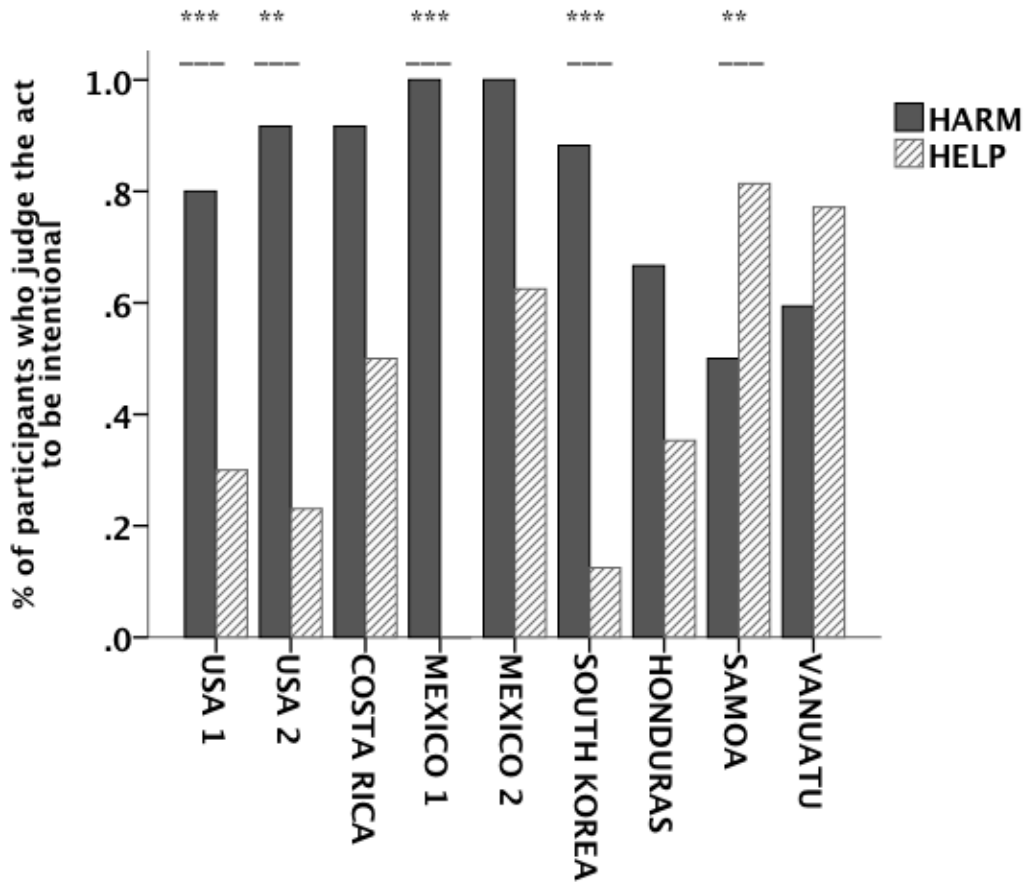


Figure 1. Percentage of participants in each culture who judge the action to be intentional in the harm and help conditions. Asterisks denote significant differences between conditions based on Fisher's exact tests; *** $p < .001$, ** $p < .01$.

Regarding the Likert scale ratings of praise or blame, a univariate ANOVA factoring condition (2) and culture (9) as between-subjects factors yielded a significant interaction of condition and culture ($F(7,269) = 8.573, p < .01, \eta^2 = .182$). Participants in USA 1, USA 2, Costa Rica, Mexico 1, and South Korea gave significantly greater ratings of blame than praise (all $p < .05$ based on two-tail pairwise comparisons with Bonferroni-adjustments). In contrast, participants in all other cultures showed either no difference between these two conditions

(Mexico 2 and Honduras), or demonstrated a slightly inverted trend toward higher praiseworthiness (Figure 2).

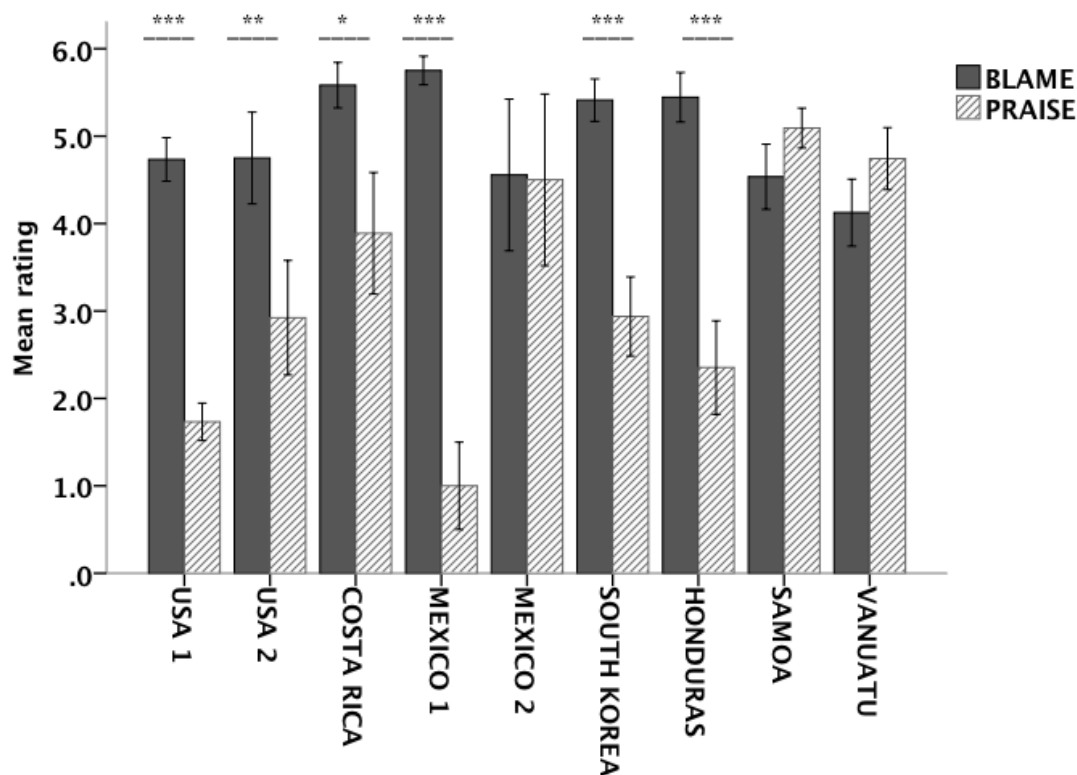


Figure 2. Mean ratings for blame (in the harm condition) and praise (in the help condition) as a function of culture. Asterisks denote significant pairwise comparisons between conditions (two-tailed with Bonferroni corrections), *** $p < .001$, ** $p < .01$, * $p < .05$. Bars represent +/- 1 SEM.

We also analyzed the association between intentional action and blame/praise ratings, following the observation in the original Knobe study (2003) these variables were strongly positively correlated. A Spearman correlation yielded a significant relationship between judgments of intentional action and praise ratings ($r_s = .520$, $N = 178$, $p < .001$). The same positive association held for judgments of intentional action and blameworthiness ratings ($r_s = .325$, $N = 166$; $p < .001$). Table 2 below reports the results of this analysis for each population. As seen in the table, the strength of the association is not equal across populations. For some populations, an

association exists only between intentional action and blame (USA 1, USA 2, Mexico 1, Mexico 2, Honduras, and South Korea; all $p < .05$), as for others no association exists at all (Costa Rica). Finally, in the two South Pacific populations, an association exists only between judgments of intentional action and praise ratings (both $p < .001$).

Table 2. Correlations between judgments of intentional action and ratings of blameworthiness (harm condition) and praiseworthiness (help condition) as a function of population.

<u>POPULATION</u>	<u>HARM</u>	<u>HELP</u>	<u>OVERALL</u>
USA 1	.658***	.343	.681***
USA 2	.463*	.248	.443*
COSTA RICA	.172	.367	.308
MEXICO 1	---- ^a	---- ^b	.907***
MEXICO 2	---- ^a	.477	.296
KOREA	.629***	.338	.643***
HONDURAS	.515*	.388	.571***
SAMOA	.252	.325***	.310**
VANUATU	.260	.529***	.412***
TOTAL	.325***	.520***	.488***

Note: All coefficients represent Spearman correlations between the intentional action and rating variables. ^aValues represent a constant; all participants judged the protagonist as intentionally harming the environment. ^bValues represent a constant; all participants judged that the protagonist did not intentionally help the environment. *** $p < .001$; ** $p < .01$, * $p < .05$.

These results partially confirm the original side-effect effect reported in the Knobe (2003a) study that was conducted with only U.S. participants. In our data, a sample of participants in Mexico 2 demonstrated no contrast between ratings of praise or blame, and

participants in Samoa and Vanuatu even trended toward an inversion of the phenomenon. Furthermore, the correlation between intentional action and praise/blame ratings is markedly different across cultures.

In a subsequent analysis we examined the relative consistency of individual responses for the subset of participants ($N= 173$) in Mexico 1, Mexico 2, Samoa, and Vanuatu who opportunistically received both help and harm vignettes. We first examined whether or not participants demonstrated the side-effect effect by changing their judgment of intentional action between help and harm conditions. Results yielded a significant effect of culture ($\chi^2 = 24.57$, $df = 3$, $p < .001$, Cramer's $V = .380$). A significant majority of participants in Mexico 1 (100%) and Mexico 2 (82.4%) changed their attributions of intentional action between the two conditions (both $p < .01$), whereas only half of Samoan (50.7%) and Ni-Vanuatu (42.4%) participants did the same. Examining only the participants who changed their responses, we next compared the proportion of participants who showed the asymmetry reported by Knobe (2003a) by judging actions as intentional in the harm but not the help condition. We observed a significant effect of population ($\chi^2 = 18.75$, $df = 3$, $p < .001$, Cramer's $V = .487$). When they changed their responses, a significant majority of participants in Mexico 1 (100%) and Mexico 2 (92.9%) judged that the protagonist intentionally hurt the environment in the harm condition and did not intentionally improve the environment in the help condition (both $p < .01$ based on two-tailed binomial tests). Only a small minority of the Samoans (27%) and Ni-Vanuatu (35.7%) participants showed this same tendency. In these two cultures, the significant majority of participants (respectively 72.6% and 64.3%) judged that the protagonist did not intentionally harm the environment in the harm condition, but that he did intentionally help the environment in

the help condition (both $p < .01$ based on two-tailed binomial tests). Note that no order effects were found for judgments of intentional action in any culture.

We also analyzed individual consistency with regard to the ratings of praise and blame. A mixed-design ANOVA factoring condition (2), population (4), and condition order (2) yielded a significant interaction of condition and population ($F(3, 162) = 23.98, p = .001, \eta^2 = .308$). There was no order effect on praise and blame ratings. In follow-up tests of this the interaction, we found that participants in Mexico 1 and Mexico 2 gave significantly higher ratings in the harm condition (both $p < .001$), whereas participants in Samoa and Vanuatu once again showed an inverted tendency to provide significantly higher ratings in the help condition (both $p < .05$). In short, participants in Mexico 1 and 2 gave significantly greater ratings of blameworthiness compared to our two South Pacific populations, who gave significantly greater ratings of praiseworthiness.

In all, the results of Experiment 1 partially replicate the original side-effect effect noted by Knobe (2003a). We confirm the findings for both judgments of intentional action and ratings of blame and praise with participants in USA 1 and USA 2, as well as Mexico 1 and South Korea. We partially replicate the phenomenon with one of the variables (i.e., with intentional action or ratings of blame/praise, but not both) in Costa Rica, Mexico 2, and Honduras. The partial replication of the intentional action question may have resulted in part from our choice to measure this outcome dichotomously, which may have undercut our ability to fully detect an effect in these small samples. Finally, we observed a tendency toward an inversion of the side-effect effect in both South Pacific samples, particularly in the significant within-subjects comparison of help and harm conditions. The within-subject analysis of participants attests to the

consistency of the original side-effect effect in Mexico 1 and Mexico 2, and demonstrates the robustness of the inverted effect in Samoa and Vanuatu. Finally, the correlation between intentional action and praise/blame ratings is markedly different across cultures.

4. Experiment 2

To control for potential differences in the meaning and title of the protagonist in the vignette used in small-scale societies (e.g., Samoa, Vanuatu, and Costa Rica), we ran a second experiment with US participants using the same labeling (e.g., High Chief) to describe the authority figure in the story. The rationale for this second experiment was that if participants were engaged in moral reasoning per se, the change in labeling of the authority figure from CEO to Chief should not affect the side-effect effect observed in the US participants of Experiment 1. Inversely, if differences in semantics drive the effect, we reasoned that the asymmetry between blame and praise conditions should either change or disappear. We expected the former.

4.1 Method

We tested an additional cohort of 56 American participants (see USA 3 in Table 1) with the English translation of the script used in these societies and that replaced CEO with High Chief (see Supplemental Materials). In a within-subjects design, participants answered both the intentionality and blame/praise rating questions for both help and harm conditions in a counterbalanced order.

4.2 Results

Following the analysis of Experiment 1, we initially considered only responses to the first vignette read by the participant (harm or help). With regard to the intention question, we

observed a significant effect of condition (Fisher's exact test: $p < .01$). Confirming the side-effect effect, 82.1% of participants in the harm condition judged the actions of the high chief as intentional, whereas only 14.3% of participants judged the action as intentional in the help condition.

Regarding the Likert scale ratings of praise or blame, an independent samples t-test yielded a significant effect of condition ($t(54) = 6.56, p < .01$, Cohen's $d = 1.79$). Consistent with the side-effect effect, ratings were significantly higher in the harm condition ($M=5.04, SD=1.07$) than the help condition ($M=2.61, SD=1.64$).

We also analyzed the within-subject consistency of the side-effect effect. A significant majority of participants ($N=40$, or 71%) changed their judgment of intentional action between conditions (two-tailed binomial test: $p = .002$). Of these participants, a significant majority ($N=38/40$, or 95%) demonstrated the side-effect effect by judging that the High Chief acted intentionally in the harm but not the help condition (two-tailed binomial test: $p < .01$). No order effects were noted. Regarding the Likert scale ratings of praise and blame, a mixed-design ANOVA factoring condition (2) as a repeated factor and condition order (2) as a between-subjects factor yielded only a significant main effect of condition ($F(1,54) = 49.10, p < .01, \eta^2 = .476$). Within-subjects, participants gave significantly higher ratings in the harm condition ($M=4.77, SD=1.07$) relative to the help condition ($M=2.61, SD=1.71$).

A final analysis demonstrates that the magnitude of the side-effect effect is comparable across the three samples of American participants. Considering only responses to the first vignette (help or harm), an independent chi-square test yielded no effect of cohort on judgments of intentional action. In the harm condition, the majority of participants in USA 1 (80.0%), USA

2 (91.7%) and USA 3 (82.1%) found the protagonist's actions to be intentional, whereas in the help condition, participants in USA 1 (30%), USA 2 (23.1%), and USA 3 (14.3%) were equally unlikely to say the protagonist acted intentionally. Overall, collapsed across cohort, a significant majority (82.9%) of US participants judged the protagonist as acting intentionally in the harm condition, whereas only 22.5% of participants judged the act as intentional in the help condition ($\chi^2 = 51.43$, $df = 1$, $p < .001$, Cramer's $V = .604$), proportions that mirror Knobe's original findings. With regard to the ratings of blame and praise, and considering only responses to the first vignette, a univariate ANOVA factoring cohort (3) and condition (2) yielded only an effect of condition ($F(1,135)=79.25$, $p < .01$, $\eta^2 = .370$). Across cohorts, participants gave significantly higher ratings in the harm condition ($M=4.85$, $SD=1.33$) than in the help condition ($M=2.30$, $SD=1.68$). These findings attest to the robustness of the side-effect effect, which appears to be equally strong across samples of U.S. participants who differ markedly in their demographic features, including age, race, and education level. Furthermore, the magnitude of the side-effect effect is comparable amongst cohorts even when changing the labeling of the authority figure in the vignette. Participants tended to provide similar ratings of blame and praise as well as judgments of intentional action whether the protagonist was depicted as a CEO or village High Chief.

Collectively, these results replicate what we found with a comparable sample of participants in Experiment 1. We thus confirm that semantic difference between CEO and high chief does not impact the moral reasoning of the participants. Both scripts elicit the side-effect effect originally demonstrated by Knobe (2003a).

5. Experiment 3

In another follow-up experiment, we examined two factors that could potentially drive the inversion of the side-effect effect noted in the South Pacific. As a working hypothesis, and to explore further the cross-cultural variability found in Experiment 1, we reasoned that because these two small-scale societies emphasize deference to authority and are characterized by highly visible and transparent delegations of power (O'Meara, 1990; Triandis & Gelfand, 1982), attributions of intentional action as well as blame and praise ratings would depend on the status of the protagonist. In particular, we anticipated that for these two cultures, attributions of intentional action, as well as ratings of praise and blame, should be higher in scenarios depicting the actions of a commoner versus a chief. The additional manipulation of gender was exploratory, probing any effect of the differential status of women in Vanuatu, where gender roles are more segregated (Rivers, 1914), compared to Samoa, where women are more active in the Samoan chief (*matai*) system of governance (Shore, 1982).

5.1 Method

Participants (N=61) were tested on both the original harm and help vignettes, as well as a second set of vignettes featuring protagonists varying in gender and status. Specifically, a subset of the Samoan (N=27) and Ni-Vanuatu (N=34) participants who read both vignettes (harm and help) about the chief in Experiment 1, also read two more help and harm vignettes involving a protagonist who was either a male or female commoner. Therefore, in this mixed-design, participants received in a counterbalanced order a set of two vignettes (help and harm) about the High Chief, plus a set of two vignettes (help and harm) about a commoner (i.e., an individual that was not a Chief) identified as either a man or a woman (see Supplemental Materials). For

our dependent measures, for each vignette participants made the same judgments of intentional action and ratings of blame or praise described previously.

5.2 Results

In a first set of analyses, we assessed whether judgments of intentional action depended on the gender of the story protagonist. A chi-square test of independence compared the proportion of participants who judged the act to be intentional as a function of culture, condition (help and harm), and protagonist gender, yielding no significant results. Across culture and conditions, participants were at chance in judging whether a woman (62.6%) or a man (66.7%) acted intentionally. The gender of the commoner did not appear to influence determinations of intentional action in either condition for either culture.

We also analyzed the Likert ratings of blameworthiness and praiseworthiness in a mixed-design ANOVA including condition (help and harm) as a within-subjects factor and culture (2) and gender (2) of the story protagonist as between-subjects factors. Results yielded only a marginal main effect of condition that is consistent with our previous findings ($F(1,57) = 3.25$, $p = .077$, $\eta^2 = .054$). Independent of culture and the gender of the story protagonist, participants tended to give higher praise ratings in the help condition ($M = 4.79$, $SD = 2.00$) than blame ratings in the harm condition ($M = 3.95$, $SD = 2.48$). Since we observed no effect of gender on either of our dependent measures, responses for the male and female vignettes are subsequently collapsed for all analyses reported below.

Next we assessed whether participants judged intentional actions differently for a Chief versus a commoner (independent of gender). A chi-square test of independence compared the proportion of participants who judged the act as intentional as a function of population, condition, and status. Results yielded a significant joint influence of these three factors ($\chi^2 =$

28.62, $df=3$, $p<.01$, Cramer's $V = .345$). In follow-up tests, for each culture we compared the proportion of participants who judged the actions as intentional as a function of condition and status. We observed that in both populations, in the harm condition participants were equally likely to attribute intentional action to the chief or the commoner. However, in the help condition, significantly more participants in both Samoa and Vanuatu judged that the Chief acted intentionally (92.3% and 88.6%, respectively) compared to the commoner (42.3% and 55.9%, respectively; both $p<.01$ based on Fisher's exact tests). These results replicate the findings of Experiment 1, but only when the protagonist in the story is in a position of authority and leadership (i.e., a chief).

Finally, we examined the rating responses in a mixed-design ANOVA including condition (help and harm) and status (chief versus commoner) as within-subjects factors and population as a between-subjects factor. Results yielded a significant interaction of condition and status ($F(1,55) = 3.51$, $p = .047$, $\eta^2 = .060$). In follow-up tests, across cultures we found that in the harm condition ($p = .008$), participants gave significantly higher ratings to the commoner ($M=5.08$, $SD=1.52$) than the chief ($M=3.84$, $SD=2.46$), indicating that they found the actions of the commoner to be significantly more blameworthy. In contrast, in the help condition, participants across culture gave significantly higher ratings to the chief ($M =5.04$, $SD=1.81$) than the commoner ($M=4.81$, $SD=2.03$), indicating that they found the actions of the chief to be significantly more praiseworthy ($p=.034$).

Consistent with our hypotheses, the results of Experiment 3 suggest that the inversion of the side-effect effect observed in the South Pacific may be driven by perceived status of the protagonist. We observed more judgments of intentional action and greater ratings of praiseworthiness in the help condition, but only when the protagonist was described as a chief.

In contrast, in the harm condition judgments of intentional action did not vary across protagonist status, although Likert ratings of blameworthiness were higher for commoners compared to the chief. In sum, the status (but not gender) of the protagonist appears to be a significant factor in both judgments of intentional action and moral evaluation in these two rural, traditional cultures.

6. Experiment 4

To ascertain further the role of perceived status as determinant of the side-effect effect, in a final experiment, we tested a new cohort of participants (USA 4) on both the harm and help vignettes, this time changing the protagonist's label from CEO to *co-worker*. The aim was to reproduce in the USA with college students what we manipulated in both Samoa and Vanuatu comparing vignettes involving either a commoner or a High Chief. We expected that if status is indeed a determinant factor, we should find that describing the protagonist as co-worker instead of CEO in the USA should have a mitigated effect on the side effect-effect. In short, despite the reported cultural variations, we expected to find some influence of status in the USA regarding attribution of intent and ratings of either blame or praise.

6.1 Method

Participants (N=60, see Table 1) read a modified version of both the original harm and help vignettes, presented in a counterbalanced fashion across participants, in which the protagonist was described as a co-worker in the company, implying that they were not in a position of power. Gender neutral language was used to describe the characters in the vignette. For our dependent measures and for each vignette, participants made the same judgments of intentional action and ratings of blame or praise described previously.

6.2 Results

A Fisher's exact test comparing the proportion of participants who judged the act to be intentional as a function of condition (help and harm) yielded significant results that confirm the original side effect-effect ($p < .001$). Overall, 15% of participants judged the co-worker's action to be intentional in the help condition, compared to 53.3% in the harm condition. The same significant asymmetry was found in ratings of praise or blame based on a paired samples t-test ($t(1,59) = 7.09, p < .001$, Cohen's $d = 1.85$). On average, the magnitude of praise ratings ($M = 2.28, SD = 1.57$) tended to be significantly less than blame ratings ($M = 4.32, SD = 1.55$). Within subjects comparisons show that a significant majority of participants demonstrated such an asymmetry (93%, binomial test $p < .001$). We observed no effect of order on either of these trends.

To further examine status as a factor, in a second set of analyses, we compared these data with those obtained from the USA 2 cohort who were the most similar in terms of demographics and who, in a between-subjects design, received either the help or harm vignettes with a CEO rather than a co-worker as protagonist. In this analysis, to ensure greater comparability between cohorts, we considered only the first vignette USA 4 participants received (either blame or praise). Regarding attributions of intention, a Fisher's exact test yielded a significant effect of protagonist status (CEO versus co-worker) for the harm ($p = .017$) but not the help condition. In the help condition, a minority of participants in both cohorts judged the actions of the protagonist as intentional, no matter their status (23.1% for USA 2 and 22.2% for USA 4, respectively). However, in the Harm condition, when the protagonist is portrayed as a CEO (USA 2 cohort), 91.7% of participants judged the protagonist as acting intentionally. In contrast, when described as a co-worker (USA 4 cohort), 51.5% of the participants judged the action as intentional.

However, the effect of protagonist status did not extend to the ratings of blame and praiseworthiness. Results failed to yield a significant condition-by-status interaction, and only a main effect of condition (help versus harm) was noted ($F(1,81)=20.37, p<.001, \eta^2 = .201$). Independent of status, participants gave significantly higher ratings in harm ($M = 4.51, SD = 1.63$) than in praise ($M = 2.47, SD = 2.06$) conditions. Therefore, the effect of status is not evident in relation to ratings of blameworthiness or praiseworthiness.

Overall, the results of Experiment 4 show that amongst US participants, changing the status of the protagonist tends to mitigate the side-effect effect, although not cancelling it. As hypothesized, we found that significantly more participants tended to attribute intentionality to harm the environment when the protagonist is described as CEO as opposed to co-worker, here construed as having less power status in enacting decisions within the portrayed institution.

7. Discussion

In small-scale, collectivist, and “face-to-face” societies (Fiske et al., 1998; Greenfield et al., 2003), high-ranking status arguably carries different meanings compared to the high-ranking status of CEOs in urban industrial and more individualistic (W.E.I.R.D.) societies. Blaming a high chief has different consequences in the former, as determinations of blame often carry with them the implicit consequence of punishment or correction. As noted by Hofstede, Hofstede, & Minkov (2010), individuals of societies like Samoa and Vanuatu that are highly allocentric and collectivistic have significantly less autonomy in enacting authoritative decisions (e.g., punishing a transgressor), an action that is more typically the prerogative of high ranking Chiefs. This could explain our findings of an inversion of the side-effect effect in Samoa and Vanuatu, whereby individuals tend to see more intentional action in the help condition and provide higher

ratings of praiseworthiness. The fact that such an inversion does not extend to vignettes in which the protagonist is a commoner provides additional direct support to this interpretation.

Noteworthy and further validating these findings is the fact that in a previous field trip research at these remote locations a few years prior, we tested over a hundred participants on the Knobe script with a high chief as protagonist and found the same inverted tendency toward higher ratings of praise. We could not include this previous body of observations in the present analysis as we accidentally omitted asking the intentional action judgment question. Based on the collected praise/blame ratings, this original testing confirms what is presented here, likely to be a real phenomenon and not an experimental fluke. Furthermore, it is also doubtful that the reported cross-cultural variability is due to semantic differences considering that Samoan and Ni-Vanuatu languages are profoundly different. If it were just a semantic by-product in these small-scale societies, the difference found with commoners in both would be unlikely. Finally, a differential understanding of harm to the environment across culture cannot account for our results since we do not observe an inversion of the side-effect effect in the two Mexican cohorts or Costa Rican sample that all depend on natural resources and agriculture for their survival (Eakin et al., 2014; Jurjonas et al., 2016; Moguel & Toledo, 1999; Cusack & Dixon, 2006).

The results of Experiment 2 further dismiss the possibility that semantic differences in the labeling of the authority figure drive the cultural variations of the side-effect effect observed here. The reported findings thus are likely linked to cultural differences in reasoning about intentional action and not to experimental or semantic artifacts. However, in relation to our interpretation, there is one caveat in the fact that in Costa Rica, where participants were also from a small-scale chief system society, we did not find the inversion of the phenomenon observed in both Vanuatu and Samoa. In addition to the small sample size at this locale, one

should note that in Costa Rica, there is considerable influence and economic aid coming from a stable and highly Westernized central government in the form of agroforestry, commercial food production, schooling, medical aid, and ecotourism (Dahlquist et al., 2007; Polidoro et al., 2008; Cusack & Dixon, 2006), all of which directly impact how the Bribri population manages their land as part of assigned indigenous reservations. In contrast, although subsistence crops are also linked to the well-being, livelihood, and cash economy of many regions in the South Pacific (O'Meara, 1990; Shore, 1982), much of the economic aid to Samoa and Vanuatu comes in the form of remittance inflows and foreign aid targeted toward infrastructure development (Allen, 2000; Jayaraman & Ward, 2006; Kumar, Naidu, & Kumar, 2011; McGregor et al., 2009).

The cultural variations in the side-effect effect reported here are also reinforced by the fact that evidence of such variations are based on both between and a within-subject comparisons for the subset of participants in Samoa, Vanuatu, Mexico 1, and Mexico 2 (Experiment 1), as well as USA 3 (Experiment 2). A similar mixed-design was used by Nichols & Ulatowski (2007) to assess the robustness of the side-effect effect, with data demonstrating that the asymmetry between greater ratings and judgments of intentionality for blame versus praise holds within individual participants. Here, however, we confirmed that the inversion of this side-effect effect in Samoa and Vanuatu exists in both the between and within subject comparison. Once again, this testifies to the robustness of our reported findings.

Finally, we found further support that perceived status of the protagonist can mitigate the side-effect effect across all the tested populations. Experiment 2 and 4 yield some evidence of differential attribution of intentionality depending on the status of the protagonist (High Chief and CEO versus commoner and co-worker). Consistent with recent findings that religious beliefs

attenuate the side-effect effect (Clark et al., 2017), our data suggest that cultural differences in motivation to blame or praise may influence judgments of intentional action.

However, although status seems to play an important role in participants' judgments of intentionality, it is not clear what status might mean across such highly contrasted cultures (USA versus Samoa and Vanuatu). Future research is needed to specify the extent to which participants across cultures share the same stereotypes regarding status of authority and leadership within their group. In particular, it is possible that the inversion of the side-effect effect reported here might primarily rest on stereotypes of authority and leadership in decision making that are unique to those small scale traditional Melanesian and Polynesian cultures. For example, in Vanuatu and Samoa there are particularly elaborated rituals of gifting and systems of reciprocity through which individuals may accrue influence, prestige, and transcendence (O'Meara, 1990; Shore, 1982; Vienne, 1984). It is indeed possible that in those cultures what participants are primarily basing their judgments on is a different perception of power rather than a fundamental difference in construing the relation between intentional action and moral evaluation, as it is often presumed in the side-effect effect. Likewise, the mitigating effect of status found with the US sample of Experiment 4 could also rest on the stereotype of CEOs harboring anti-environmental attitudes (see for example Sripada & Konrath, 2011; Sripada, 2012). In addition, chieftain systems of governance in the South Pacific are based on traditions that are strongly cohesive and multigenerational. There may therefore be a degree of shared understanding and consensus about leadership in these societies that is qualitatively different than the Western context of high market integration and large corporations with broader distribution of leadership making it harder to delineate authority relationships.

Inversions of the side-effect effect have been previously linked to contextual factors such as the explicitness of the moral norms that are violated in the vignette (Robinson, Stey, & Alfano, 2015), or whether the participant is an agent or observer of the unintended consequence (Feltz, Harris, & Perez, 2012). Based on our data, it is difficult to determine the extent to which participants in the South Pacific may have aligned themselves with the story protagonist, and whether thinking of themselves as beneficiaries of the protagonist's action might have made them more inclined to attribute intention to the chief's action in the help condition. However, if something about living in small-scale, collective, traditional, and/or rural societies lent itself to such projective perspective-taking or immersive thinking, we might also expect to observe similar effects in the Costa Rican or even Mexican populations, which is not the case. Further research could help disambiguate these possibilities. Finally, we do not believe that potential cultural differences in understanding of intentional action can fully explain our results. The classic side-effect effect is observed across a wide array of theory of mind concepts, including several variants of *intent* and *belief* (McCann, 2005; Beebe & Buckwalter, 2010; Guglielmo & Malle, 2010). To our knowledge, none of these articulations of desire or belief states evoke the inversion of the side-effect effect noted here. In summary, although our results do not account directly for the mechanisms linking intentional attribution and moral judgments of praise or blame, they show that factoring status can help dissociate these two dimensions and thus could help in figuring out those mechanisms in future research.

8. Conclusions

Using the original Knobe paradigm, we find an inversion of the side-effect effect in two small-scale, traditional societies of the South Pacific. Our findings suggest that the widely

documented asymmetry between judgments of intentional action and moral evaluations is not necessarily unidirectional. Cultural beliefs about moral actions may influence how people perceive unintentional consequences, with both praise and blame influencing judgments of intentional action (Clark et al., 2017).

Our results raise a number of interesting questions. If the relationship between moral evaluation and judgments of intentional action can be multi-directional, as our findings suggest, how can we predict what circumstances will increase or decrease blame attributions? Alicke et al. (2008) point to a “culpable control” model in which participants evaluate causal information alongside the favorability of the outcome or actors. Assignations of blame may be made when an agent is perceived as controlling a negative outcome (as in the harm condition), but also if evidential information suggests that an actor is dislikeable or is shown to be negligent (as might be the perception following the protagonist’s casual disregard for the environment). On the other hand, if the information about the actors or consequences is positive enough, it may even override contradictory evidence (e.g., a thriving economy or environment may be enough to deemphasize the cavalier attitude of the protagonist). Culture may influence judgments of intentional action not only by defining what acts are considered helpful versus harmful, but also perhaps by setting the threshold for how negative outcomes or agents must be before they are considered blameworthy (Alicke et al., 2008). Our findings may also call to question the assumption that transculturally, punitive measures carry higher instrumental value than reward (Henrich et al., 2006). That some cultures differentially emphasize praise over punishment may be consistent with an emerging literature demonstrating cultural variation in the tendency to engage in punitive versus restorative justice following violation of a fairness norm (Robbins & Rochat, 2011). Understanding the relationship between moral evaluation and judgment of

intentional action may have wide-reaching implications for how we understand and reform our criminal justice systems.

In sum, we interpret our data to suggest that although the tendency to engage in moral evaluations is ubiquitous—we all engage in judgments about other people—rationalization for these judgments is likely context specific. More research is needed to specify further what cultural factors influence such evaluations. The cultural differences noted here point to the central importance of taking context seriously when theorizing about moral reasoning.

Acknowledgements

Our gratitude goes to lab coordinator Theresa Nettles as well as the research assistants of the Emory University Infant and Child Lab who assisted with data collection, including Yoana Villareal in Mexico, Mariana Ruiz in Costa Rica, Eun Jie Lee in South Korea and Honduras, Davido Molho in Samoa and Vanuatu, and Jordan Boxer in New York City. We also extend our deep appreciation and thanks to the families and leaders of the communities we visited, including Chief Tofa Tuileutu and Valeua Faga in Samoa and Chief Edgar Howard in Vanuatu.

References

- Alicke, M.D., Davis, T.L., Buckingham, J.T., & Zell, E. (2008). Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin*, 34, 1371-1381.
- Allen, M. G. (2000, June). Subsistence or cash cropping? Food security on Malo island, Vanuatu. In *Aciar Proceedings* (pp. 100-111).
- Alfano, M, Beebe, J.R., & Robinson, B. (2012). The centrality of belief and reflection in Knobe-effect cases: A unified account of the data. *The Monist*, 95, 264-289.
- Beebe, J.R. (2013). A Knobe effect for belief ascriptions. *Review of Philosophy and Psychology*, 4, 235-268.
- Beebe, J.R. & Buckwalter, W. (2010). The epistemic side-effect effect. *Mind and Language*, 25, 474-498.
- Clark, C.J., Bauman, C.W., Kamble, S.V., & Knowles, E.D. (2017). Intentional sin and accidental virtue? Cultural differences in moral systems influence perceived intentionality. *Social Psychological and Personality Science*, 8, 74-82.
- Cova, F. & Naar, H. (2012). Side-effect effect without the side effects: The pervasive impact of moral considerations on judgments of intentionality. *Philosophical Psychology*, 25, 837-854.
- Cusack, D., & Dixon, L. (2006). Community-based ecotourism and sustainability: Cases in Bocas del Toro Province, Panama and Talamanca, Costa Rica. *Journal of Sustainable Forestry*, 22, 157-182.
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108, 353-380.

- Cushman, F., Sheketoff, R., Wahrton, S. & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, 127, 6-21.
- Cushman, F. & Mele, A. (2008). Intentional action: Two-and-a-half folk concepts? In J. Knobe & S. Nichols (Eds.), *Experimental Philosophy* (pp 171-188). Oxford: Oxford University Press.
- Dahlquist, R.M., Whelan, M.P., Winowiecki, L., Polidoro, B., Candela, S., Harvey, C.A., Wulfhorst, J.D., McDaniel, P.A., & Bosque-Pérez, N.A. (2007). Incorporating livelihoods in biodiversity conservation: A case study of cacao agroforestry systems in Talamanca, Costa Rica. *Biodiversity Conservation*, 16, 2311-2333.
- Ditto, P.H., Pizarro, D.A., & Tannenbaum, D. (2009). Motivated moral reasoning. In B. Ross (ed.) *Psychology of Learning and Motivation, Vol. 50* (pp. 307-338). New York: Academic Press.
- Eakin, H., Tucker, C.M., Castellanos, E., Diaz-Porras, R., Barrera, J.F., & Morales, H. (2014). Adaptation in a multi-stressor environment : Perceptions and responses to climactic and economic risks by coffee growers in Mesoamerica. *Environment, Development and Sustainability*, 16, 123-139.
- Feltz, A., Harris, M., & Perez, A. (2012). Perspective in intentional action attribution. *Philosophical Psychology*, 25, 673-687.
- Fiske, A.P., Kitayama, S., Markus, H., & Nisbett, D. (1998). The cultural matrix of social psychology. In *Handbook of Social Psychology, 4th Ed*, D. Gilbert, S. Fiske, & G. Lindzey (Eds.). New York: McGraw-Hill, pp. 915-981.
- Fletcher, G.P. (1998). *Basic Concepts of Criminal Law*. New York: Oxford University Press.

- Green, S.P. (2000). The universal grammar of criminal law. *Michigan Law Review*, 98, 2104-2125.
- Greenfield, P.M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology*, 54, 461-490.
- Guglielmo, S. & Malle, B.F. (2010). Can unintended side effects be intentional? Resolving a controversy over intentionality and morality. *Personality and Social Psychology Bulletin*, 36, 1635-1647.
- Hamlin, J.K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, 128, 451-474.
- Hamlin, J.K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450, 557-559.
- Henrich, J., Heine, S.J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61-135.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., and Ziker, J. (2006). Costly punishment across human societies. *Science* 312, 1767–1770.
- Hofstede, G., Hofstede, G. J., Minkov, M. (2010). *Cultures and Organizations: Software of the Mind*, 3rd ed. New York: McGraw-Hill.
- Jayaraman, T.K. & Ward, B.D. (2006). Economic growth in a vulnerable island nation: An empirical study of the aid-growth nexus in Vanuatu. *Asia-Pacific Development Journal*, 13, 93-112.

- Jurjonas, M., Crossman, K., Solomon, J., & Baez, W.L. (2016). Potential links between certified organic coffee and deforestation in a protected area in Chiapas, Mexico. *World Development*, 78, 13-21.
- Knobe, J. (2003a). Intentional action and side effects in ordinary language. *Analysis*, 63, 190-193.
- Knobe, J. (2003b). Intentional action in folk psychology: An empirical investigation. *Philosophical Psychology*, 16, 309-323.
- Knobe, J. (2005). Theory of mind and moral cognition: Exploring the connections. *Trends in Cognitive Sciences*, 9, 357-359.
- Knobe, J. (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies*, 130, 203-231.
- Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, 33, 315-365.
- Knobe, J. & Burra, A. (2006). The folk concepts of intention and intentional action: A cross-cultural study. *Journal of Cognition and Culture*, 6, 113-132.
- Knobe, J. & Mendlow, G. (2004). The good, the bad, and the blameworthy: Understanding the role of evaluative reasoning in folk psychology. *Journal of Theoretical and Philosophical Psychology*, 24, 252-258.
- Kumar, R. R., Naidu, V., & Kumar, R. (2011). Exploring the nexus between trade, visitor arrivals, remittances and income in the Pacific: a study of Vanuatu. *Acta Universitatis Danubius. Œconomica*, 7(4), 199-218.
- Lagnado, D. A. & Channon, S. (2008) Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, 108, 754-770.

- Lanteri, A. (2012). Three-and-a-half folk concepts of intentional action. *Philosophical Studies*, 158, 17-30.
- Lesle, A. Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, 17, 421-427.
- McGregor, A., Bourke, R.M., Manley, M., Tubuna, S., & Deo, R. (2009). Pacific island food security : Situation, challenges, and opportunities. *Pacific Economic Bulletin*, 24, 24-42.
- Malle, B.F., Moses, L.J., & Baldwin, D.A. (2001). Introduction: The significance of intentionality. In B.F. Malle, L.J. Moses, & D.A. Baldwin (Eds.), *Intentions and Intentionality: Foundations of Social Cognition*. Cambridge, MA: MIT Press.
- Mallon, R. (2008). Knobe vs. Machery: Testing the trade-off hypothesis. *Mind and Language*, 23, 247-255.
- Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Science*, 11, 143-152.
- Moguel, P. & Toledo, V.M. (1999). Biodiversity conservation in traditional coffee systems of Mexico. *Conservation Biology*, 13, 11-21.
- Nadelhoffer, T. (2004a). Blame, badness, and intentional action: A reply to Knobe and Mendlow. *Journal of Theoretical and Philosophical Psychology*, 24, 259-269.
- Nadelhoffer, T. (2004b). On praise, side effects, and folk ascriptions of intentionality. *Journal of Theoretical and Philosophical Psychology*, 24, 196-213.
- Nadelhoffer, T. (2005). Skill, luck, control, and intentional action. *Philosophical Psychology*, 18, 341-352.
- Nadelhoffer, T. (2006). Desire, foresight, intentions, and intentional actions: Probing folk intuitions. *Journal of Cognition and Culture*, 6, 133-157.

- Nadelhoffer, T. (2011). Experimental philosophy of action. In J. Aguilar, A. Buckareff, & K. Frankish (Eds.), *New Waves in the Philosophy of Action*. New York: Palgrave-MacMillan.
- Nichols, S. & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind and Language*, 22, 346-365.
- Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention, and outcome on children's moral judgments. *Journal of Experimental Child Psychology*, 104, 382-397.
- O'Meara, T. (1990). *Samoan Planters: Tradition and Economic Development in Polynesia*. Fort Worth: Holt, Rinehart, & Winston, Inc.
- Pellizzoni, S., Girotto, V., & Surian, L. (2010). Beliefs and moral valence affect intentionality attributions: The case of side effects. *Review of Philosophy and Psychology*, 1, 201-209.
- Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind and Language*, 24, 586-604.
- Polidoro, B.A., Dahlquist, R.M., Castillo, L.E., Morra, M.J., Somarriba, E., & Bosque-Pérez, N.A. (2008). Pesticide application practices, pest knowledge, and cost-benefits of plantain production in the Bribri-Cabécar Indigenous Territories, Costa Rica. *Environmental Research*, 108, 98-106.
- Rivers, W.H.R. (1914). *The History of Melanesian Society*. Cambridge University Press.
- Robbins, E., & Rochat, P. (2011). Emerging signs of strong reciprocity in human ontogeny. *Frontiers in Psychology*, 2 (353), 1-14.
- Robinson, B., Stey, P., & Alfano, M., (2015). Reversing the side-effect effect: The power of salient norms. *Philosophical Studies*, 172, 177-206.

- Shepard, J. & Wolff, P. (2013) Intentionality, evaluative judgments, and causal structure. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3390-3395). Berlin: Cognitive Science Society.
- Shore, B. (1982). *Sala'ilua: A Samoan Mystery*. New York: Columbia University Press.
- Sousa, P. & Holbrook, C. (2010). Folk concepts of intentional action in the contexts of amoral and immoral luck. *Review of Philosophy and Psychology*, 1, 351-370.
- Sripada, C. & Konrath, S. (2011). Telling more than we can know about intentional action. *Mind & Language*, Vol. 26, Issue 3, pp. 353-380.
- Sripada, C. (2012), Mental state attribution and the side-effect-effect. *Journal of Experimental Social Psychology*. Vol. 48, Issue 1, pp. 233-238.
- Tannenbaum, D., Ditto, P.H., & Pizarro, D.A. (2007). Different moral values produce different judgments of intentional action. Unpublished manuscript, University of California-Irvine.
- Triandis, H. C., and Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *J. Pers. Soc. Psychol.* 74, 118–128.
- Uttich, K. & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition*, 116, 87-100.
- Vienne, B. (1984). *Gens de Motlav. Idéologie et pratique sociale en Mélanésie*. Publications de la Société des Océanistes, 42. Paris: Société des Océanistes.
- Wright, J. & Bengson, J. (2009). Asymmetries in judgments of responsibility and intentional action. *Mind and Language*, 24, 24-50.

- Young, L., Cushman, F., Adolphs, R., Daniel, T., Hauser, M. (2006). Does emotion mediate the relationship between an action's moral status and its intentional status? Neuropsychological evidence. *Journal of Cognition and Culture*, 6, 291-304.
- Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, 104, 8235-8240.
- Zalla, T. & Leboyer, M. (2011). Judgment of intentionality and moral evaluation in individuals with high functioning autism. *Review of Philosophical Psychology*, 4, 681-698.
- Zalla, T., Machery, E., & Leboyer, M. (2010). Intentional action and moral judgment in Asperger syndrome and high-functioning autism. Unpublished Manuscript. Institut Jean Nicod – CNRS.