

Supplementary materials

Bence Bago

14 January 2021

```
# Show the number of participants excluded because of not finished responses
#trolley_raw %>%
# filter(str_detect(str_to_lower(practice), "false")) %>%
# filter(!(lab == "TUR_021" & StartDate < date("2020-04-22")),
#        !(lab == "AUT_003" & StartDate < date("2020-06-18")),
#        !(lab == "USA_095")) %>%
# filter(Progress < 98) %>%
# count()
```

1. Demographics and exclusion

2. Main replication analysis

Study 1a

Bayesian analysis

Table S1: The effect of personal force on moral dilemma judgements on Trolley dilemmas

Exclusion	Cluster	BF	RR	t	df	p
Exclusion	Eastern	8.0e+01	1.5e-02, 4.7e+01	-3.41	553.15	<.001
	Southern	8.0e+01	3.0e-08, 9.0e+06	-3.41	553.15	<.001
	Western	8.0e+01	4.0e-0.3, 1.4e+01	-3.41	553.15	<.001
No exclusion	Eastern	4.0e+14	8.0e-15, 2.0e+13	-8.56	1707.97	<.001
	Southern	4.0e+17	1.0e-23, 1.9e+22	-9.40	1851.54	<.001
	Western	4.5e+14	2.0e-19, 6.0e+17	-8.57	2112.23	<.001

Table S1 and S2 and Figures S1-4 summarise the results of the effect of personal force on moral dilemma judgements, separately for the Speedboat and Trolley dilemmas. First, we conducted the analysis excluding participants based on our exclusion criteria. In this analysis, we replicated the effect of personal force in all cultural clusters, in both the Trolley and Speedboat dilemmas. We also analysed the data without exclusions - in both dilemmas and we arrive at the same conclusion; the effect of personal force on moral judgments is culturally universal; people judge acts that require no personal force more acceptable than acts that do. The “RR” column in the tables indicates the robustness region; a region for the prior with which we would arrive at the same conclusion. This region is very wide, indicating that the results are robust to our prior choices. Although excluding the data of a large number of participants did not change the conclusion, it introduced some noise that differentially affected the cultural clusters. In the western cluster, for example, the t value (effect size) is smaller if we analyse the data after exclusions, while such difference cannot be observed in the two other cultural clusters.

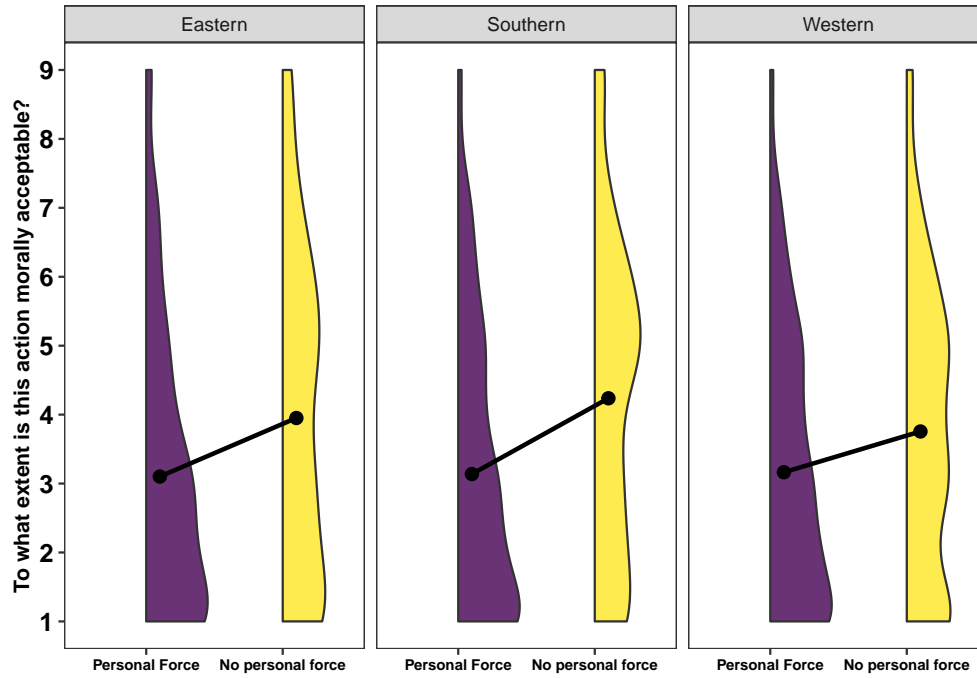


Figure S1: Results on Trolley dilemmas in Study 1 with all exclusion criteria applied. Black full circles indicate group level means.

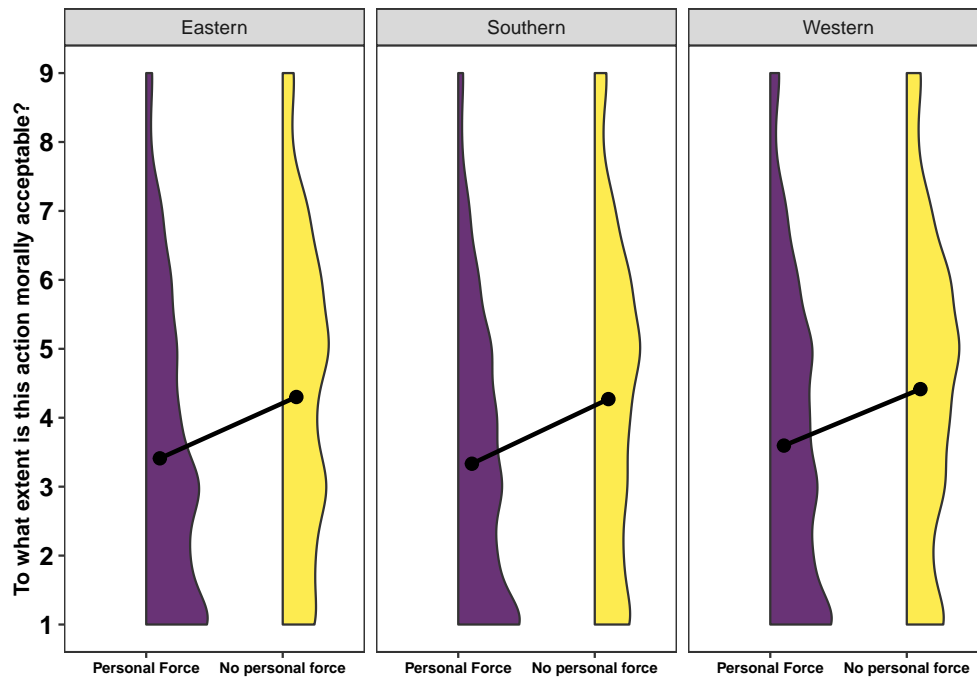


Figure S2: Results on Trolley dilemmas in Study 1 with no exclusion criteria applied. Black full circles indicate group level means.

Table S2: The effect of personal force on moral dilemma judgements on Speedboat dilemmas

Exclusion	Cluster	BF	RR	t	df	p
Exclusion	Eastern	1.2e+05	2.6e-06, 5.7e+04	-5.26	283.92	<.001
	Southern	1.0e+03	3.0e-04, 2.5e+03	-4.19	436.86	<.001
	Western	2.5e+01	6.0e-03, 4.3	-3.01	437.36	0.003
No exclusion	Eastern	1.4e+09	8.4e-10, 1.1e+08	-6.84	1708.66	<.001
	Southern	9.4e+10	5.8e-18, 2.2e+16	-7.48	1855.02	<.001
	Western	4.5e+13	8.0e-18, 1.2e+16	-8.29	2122.07	<.001

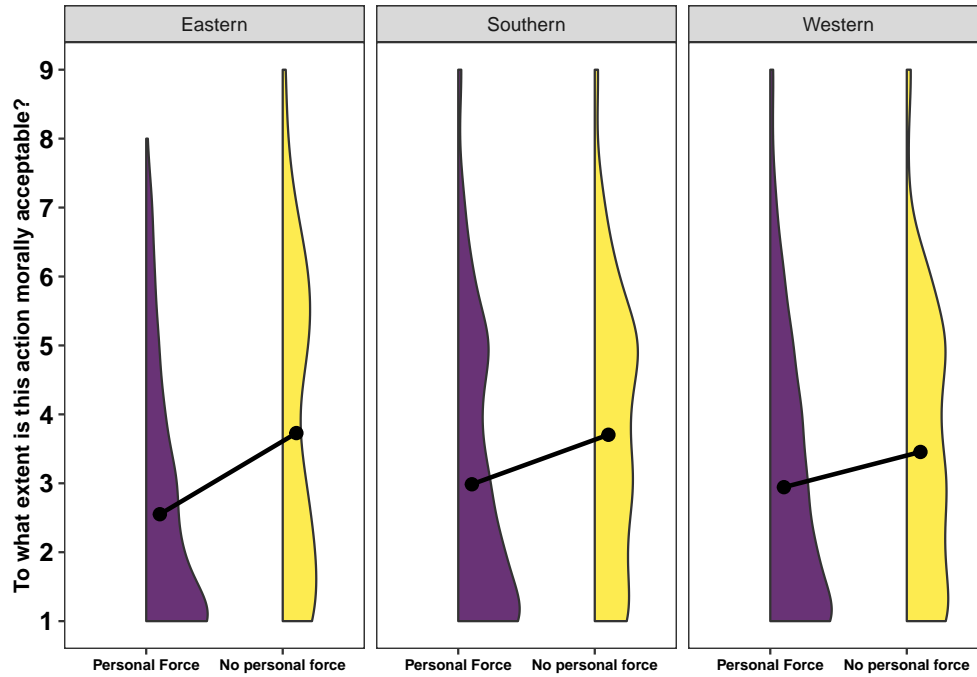


Figure S3: Results on Trolley dilemmas in Study 1 with all exclusion criteria applied. Black full circles indicate group level means.

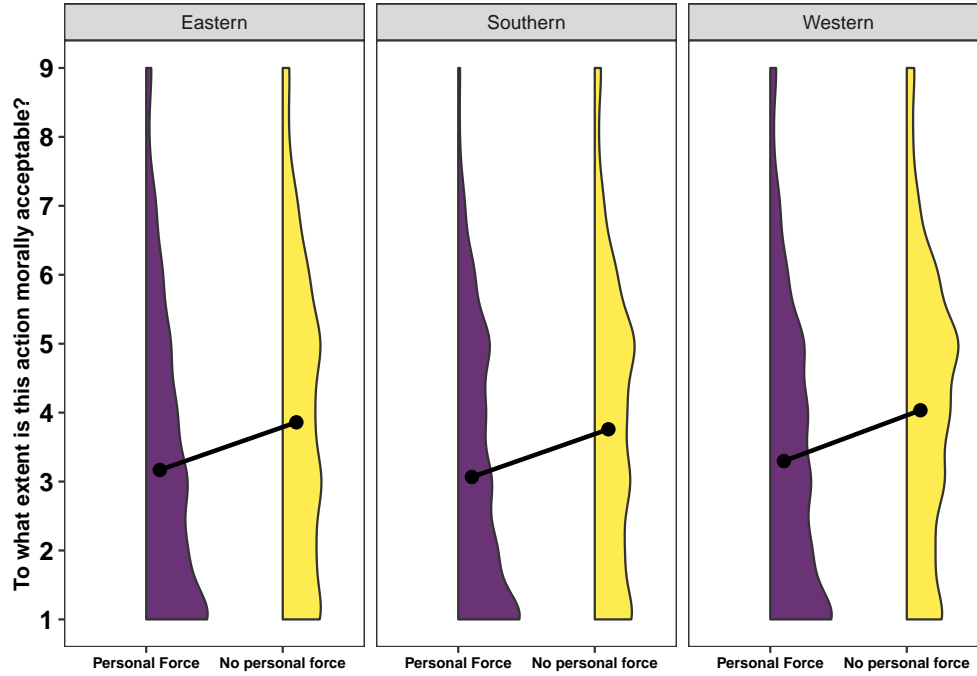


Figure S4: Results on Speedboat dilemmas in Study 1 with no exclusion criteria applied. Black full circles indicate group level means.

Study 2a and 2b

We analysed the interactional effect of personal force and intention in a similar way (Tables S3 and S4, Figures S5-S8). We treatment coded both personal force (0 = no personal force; 1 = personal force) and intention (0 = no intention; 1 = intention) and added them to the model. In the Bayesian analysis, the Bayes factor represents the ratio of two Bayesian regression models (we always used the same priors for these models): the denominator model, a model with the two main effects but without the critical two-way interaction, and in the numerator the full model including the interaction term. In case of the trolley dilemmas, we have strong evidence ($BF > 10$) to conclude that the interaction exists in the Western and Southern cultural clusters, while in the Eastern cluster results are inconclusive ($1/10 < BF < 10$). Here, we believe is again a case of selection bias. When we analysed all the data, without exclusions, we replicated the effect in all cultural clusters ($BF > 10$). The evidence however is more mixed for the Speedboat dilemmas. Here, we found evidence only for the interaction on the Western cluster but we did not find an effect on the other two clusters. Although the BF goes below only our null threshold (eastern cluster - without exclusions), $BF < 1$ in this case makes the null hypothesis more likely than the alternative one.

Table S3: Do personal force interact with intention on Trolley dilemmas?

Exclusion	Cluster	BF	RR	F	df	p
Exclusion	Eastern	6.0e-01	2.2e-01, Inf	0.041	1, 319	0.84
	Southern	4.5e+00	1.3e-0.2, 2.3	7.583	1, 619	0.006
	Western	7.2e+08	<1.0e-04, 3.1e+10	47.011	1, 2551	<.001
No exclusion	Eastern	1.9e+03	2.7e-0.4, 1.1e+03	18.430	1, 2052	<.001
	Southern	1.3e+14	<1.0e-04, 8.9e+15	72.282	1, 2485	<.001
	Western	6.4e+48	<1.0e-04, >1.0e+18	233.910	1, 12874	<.001

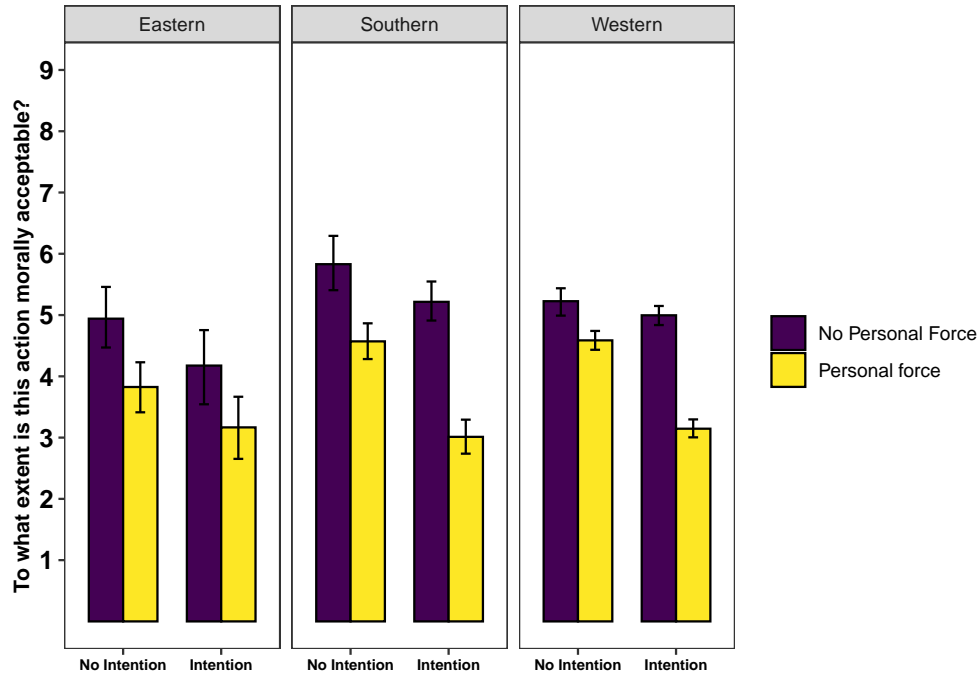


Figure S5: Results on Trolley dilemmas in Study 2 when all exclusion criteria are applied. Error bars are 95% confidence intervals on the mean.

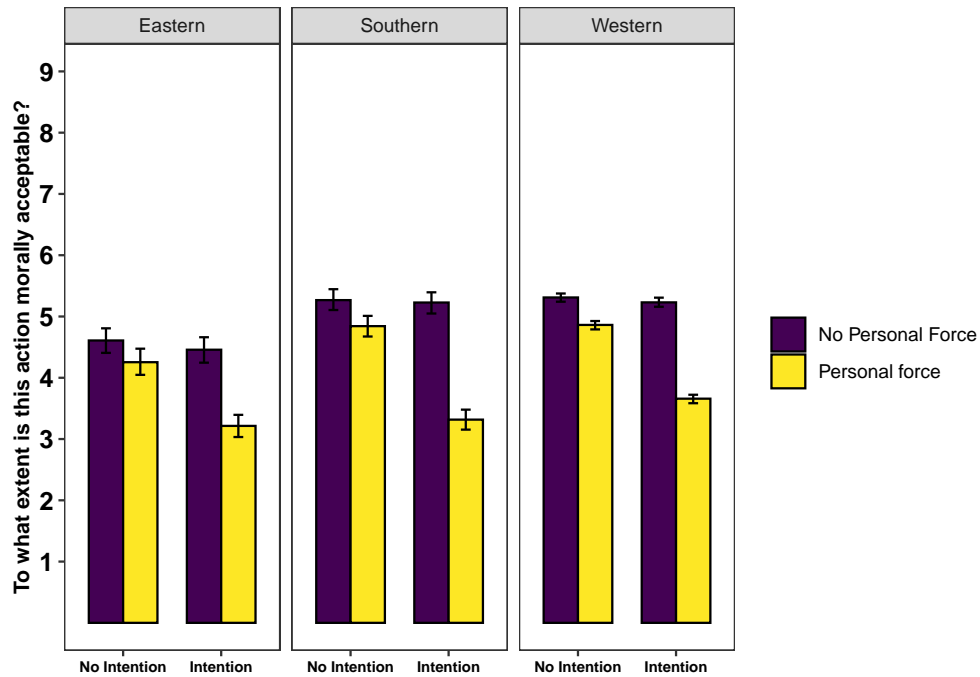


Figure S6: Results on Trolley dilemmas in Study 2 when no exclusion criteria are applied. Error bars are 95% confidence intervals on the mean.

Table S4: Do personal force interact with intention on Speedboat dilemmas?

Exclusion	Cluster	BF	RR	F	df	p
Exclusion	Eastern	0.7	0, 2.8e-01	1.283	1, 319	0.258
	Southern	0.4	2.2e-01, Inf	1.060	1, 619	0.304
	Western	0.1	1.2e-03, 1.5e+01, 3.1e+10	0.268	1, 2551	0.604
No exclusion	Eastern	0.1	8.0e-02, Inf	0.001	1, 2052	0.979
	Southern	0.2	0, 3.5e-01	1.050	1, 2485	0.306
	Western	44.0	2.0e-03, 7.8e-01	13.585	1, 12874	<.001

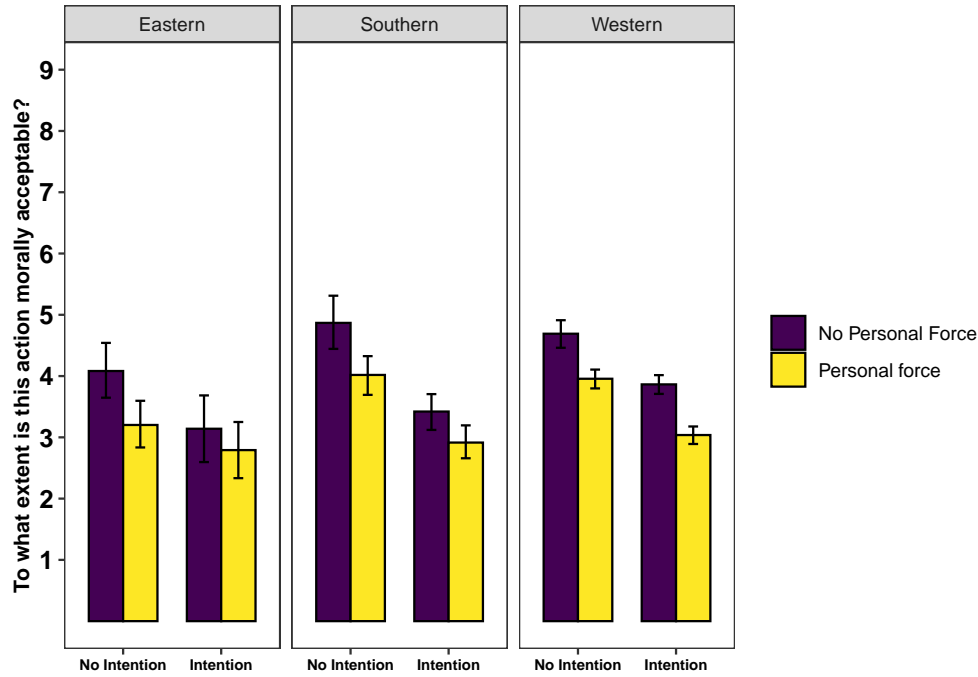


Figure S7: Results on the Speedboat dilemmas in Study 2 when all exclusion criteria are applied. Error bars are 95% confidence intervals on the mean.

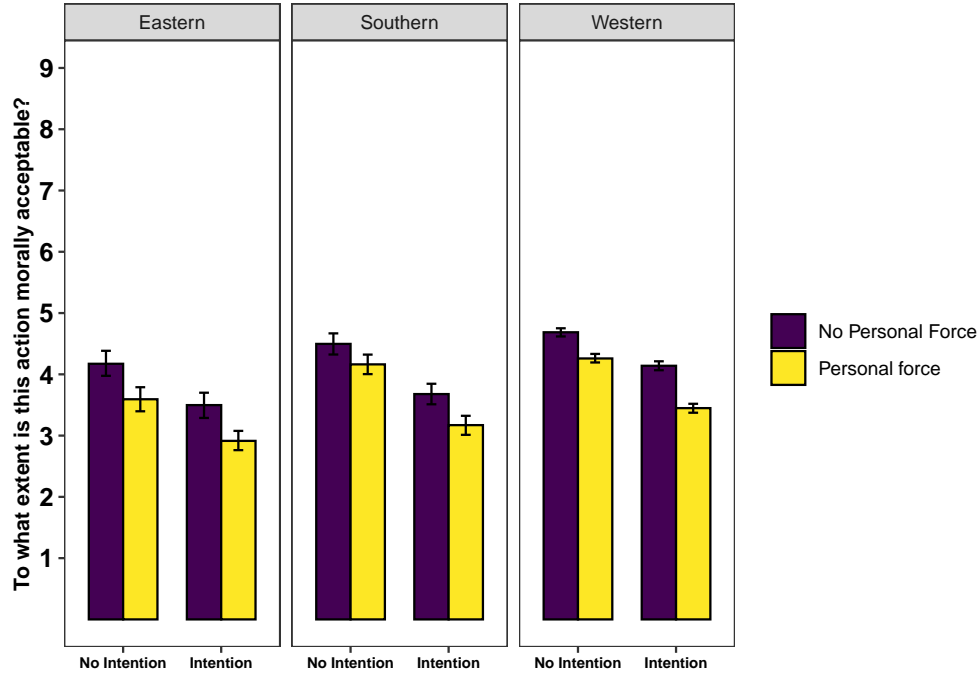


Figure S8: Results on the Speedboat dilemmas in Study 2 when no exclusion criteria is applied. Error bars are 95% confidence intervals on the mean.

The initial idea behind the exclusion criteria was that by excluding certain people, we decrease noise and, as a result, we avoid false negative effects. For example, we thought that people failing the attention check questions do not paid enough attention to the task, or that people who were familiar with the trolley problem would respond with their knowledge of the trolley problem in mind, and that would result in a decreased effect of personal force. We observed the complete opposite; we observed a bigger effect size if there was no participants excluded in all cases previously analysed. This could mean either an increased random noise due to the decreased N, or a systematic bias that was intruduced by a large number of exclusions (85% excluded overall). Hence, to avoid such biases, from this point, we applied no exclusion criteria.

3. Individualism-collectivism analysis

With all exclusions

Table S5: Is the interaction of personal force and intention affected by individualism/collectivism on Trolley dilemmas whhen all exclusion criteria is applied?

variable	BF	b	p
Country-level collectivism	0.6107	-5.632	0.301
Vertical Individualism	1.1520	-0.102	0.247
Horizontal Individualism	1.1000	0.035	0.744
Vertical Collectivism	0.9130	0.025	0.791
Horizontal Collectivism	1.3490	-0.074	0.487

Table S6: Is the interaction of personal force and intention affected by individualism/collectivism on Speedboat dilemmas when all exclusion criteria is applied?

variable	BF	b	p
Country-level collectivism	1.681	2.021	0.704
Vertical Individualism	1.930	-0.130	0.129
Horizontal Individualism	2.033	-0.003	0.980
Vertical Collectivism	1.281	0.077	0.395
Horizontal Collectivism	1.575	-0.096	0.354

Without any exclusions

We applied similar analysis strategy as before: we compared the full model (including the three way interaction) with a model that includes all main effects and interactions but the three-way interaction. Here, the results are somewhat fuzzier. IN the trolley dilemma, the frequentist test shows a significant effect of both country level and horizontal collectivism, but the Bayes Faktor shows the contrary; that the null effect is more likely to be true than the alternative. This comes from the differences between the two tests; the frequentists compares to the intercept, while the Bayesian to an alternative model. Even if the effect exists it cannot be generalized to the speedboat dilemmas.

Table S7: Is the interaction of personal force and intention affected by individualism/collectivism on Trolley dilemmas?

variable	BF	b	p
Country-level collectivism	0.9458	-4.123	0.047
Vertical Individualism	0.3432	0.006	0.881
Horizontal Individualism	2.5320	-0.086	0.064
Vertical Collectivism	0.4408	-0.009	0.827
Horizontal Collectivism	6.6530	-0.107	0.020

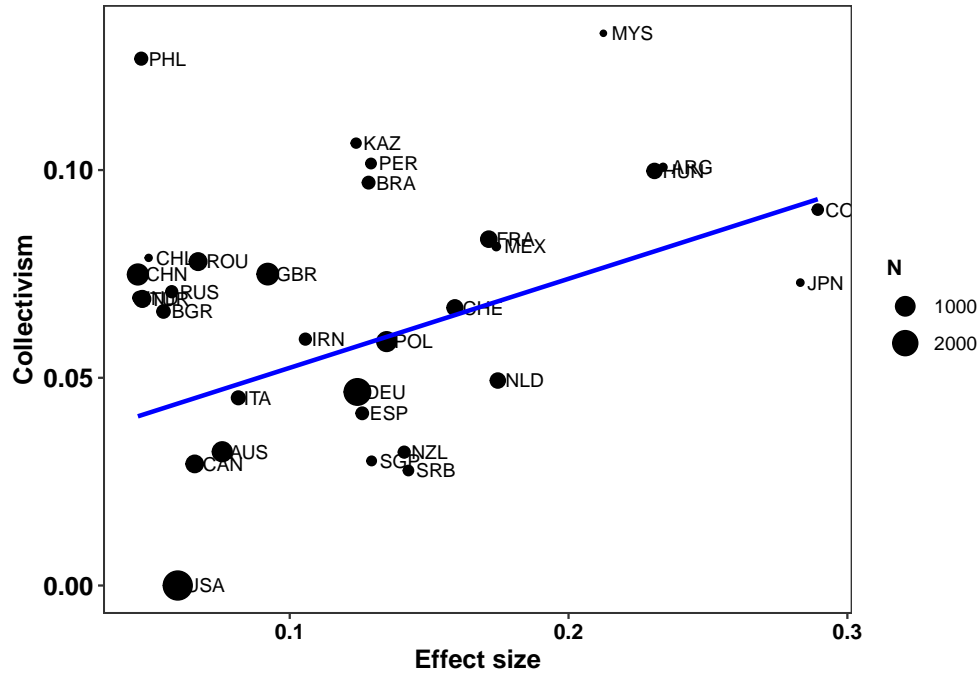


Figure S9: Correlation between country-level individualism/collectivism and the effect of personal force in Eta squared on the Trolley dilemmas

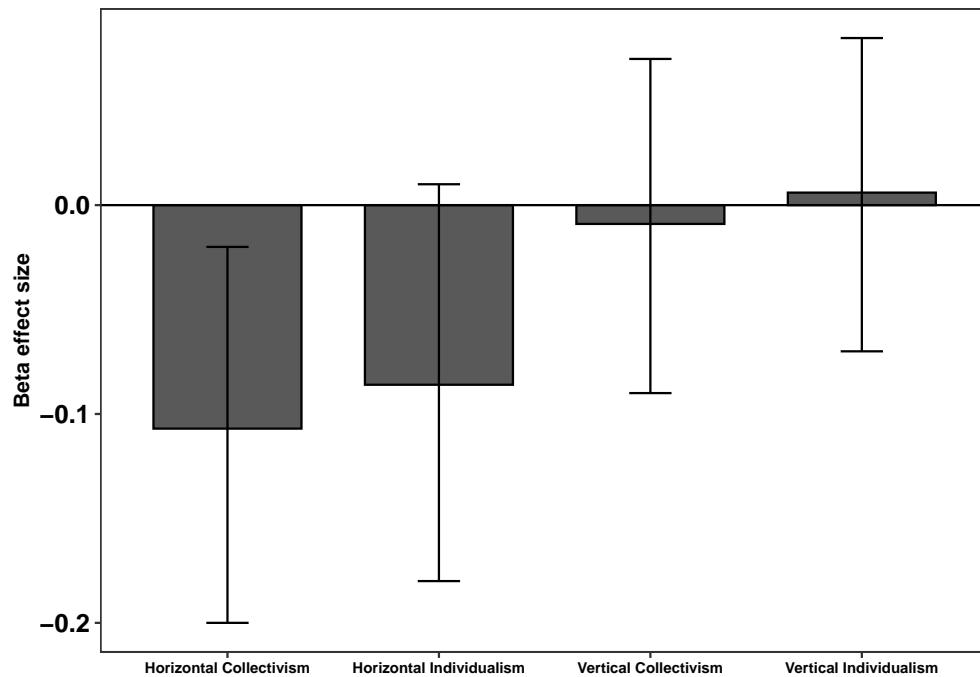


Figure S10: Personal level individualism/collectivism effects on the interaction of personal force and intention (trolley dilemmas)

Table S8: Is the interaction of personal force and intention affected by individualism/collectivism on Speedboat dilemmas?

variable	BF	b	p
Country-level collectivism	1.1480	4.420	0.027
Vertical Individualism	0.1565	0.004	0.921
Horizontal Individualism	2.1720	-0.035	0.440
Vertical Collectivism	0.5043	0.015	0.688
Horizontal Collectivism	0.4651	-0.036	0.424

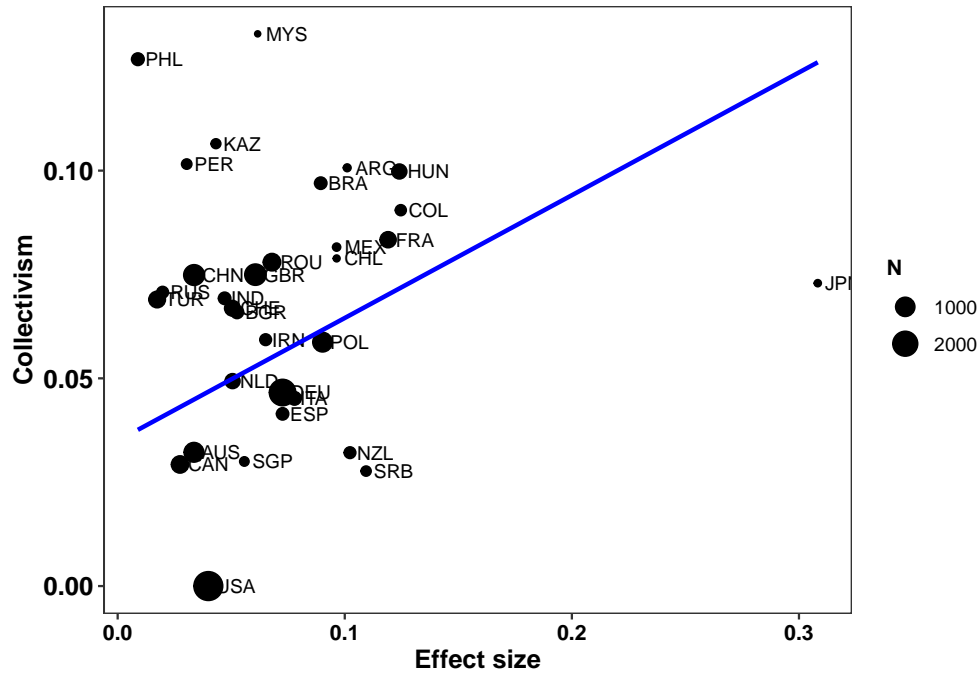


Figure S11: Correlation between country-level individualism/collectivism and the interactional effect of personal force and intention in Eta squared on Speedboat dilemmas

Table S9: Effect of Physical Contact and Intention

Cluster	Dilemma	Comparison	t	Bf	df	p
Southern	Trolley	Physical Contact	0.15	1.50e-01	1,344.04	0.88
Southern	Speedboat	Physical Contact	-1.02	2.50e-01	1,310.01	0.31
Southern	Trolley	Intention	-8.51	9.63e+13	1,293.65	0.00
Southern	Speedboat	Intention	-6.78	4.50e+08	1,279.18	0.00
Western	Trolley	Physical Contact	-0.83	1.50e-01	1,837.66	0.40
Western	Speedboat	Physical Contact	-2.12	9.50e-01	1,877.25	0.03
Western	Trolley	Intention	-11.38	1.40e+28	1,643.71	0.00
Western	Speedboat	Intention	-8.98	3.67e+16	1,690.51	0.00
Eastern	Trolley	Physical Contact	1.68	6.00e-01	1,074.99	0.09
Eastern	Speedboat	Physical Contact	2.33	1.87e+00	1,134.47	0.02
Eastern	Trolley	Intention	-2.38	2.51e+00	1,066.74	0.02
Eastern	Speedboat	Intention	-2.51	3.37e+00	1,076.70	0.01

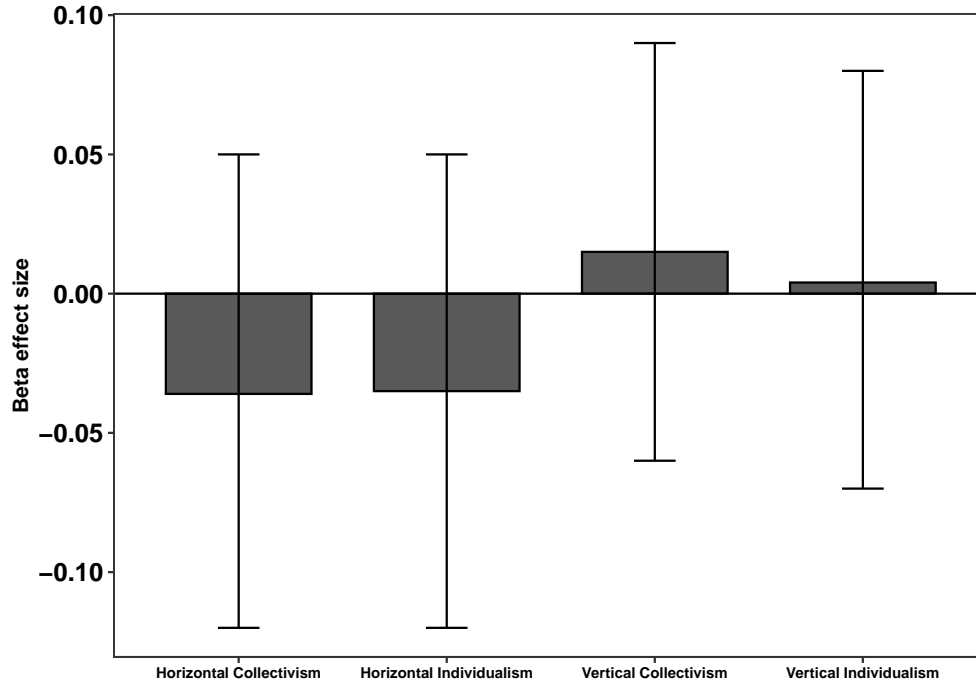


Figure S12: Personal level individualism/collectivism effects on the interaction of personal force and intention (speedboat dilemmas)

4. Additional analysis

Effect of physical contact and intention

In every cluster and for both types of dilemma we found good enough evidence supporting the alternative hypothesis when testing the effect of physical contact and the effect of intention. The summary of the results can be found in Table X.

Table S10: Comparing the Standard Switch and Standard Footbridge Dilemmas

Cluster	Dilemma	t	Bf	df	p
Southern	Trolley	15.77	3.67e+47	1,225.91	0.00
Southern	Speedboat	11.27	4.93e+24	1,234.52	0.00
Western	Trolley	31.79	6.41e+201	6,483.94	0.00
Western	Speedboat	24.84	1.12e+126	6,482.00	0.00
Eastern	Trolley	9.96	1.02e+19	1,044.83	0.00
Eastern	Speedboat	9.58	3.61e+17	1,032.24	0.00

Comparing the standard switch and standard footbridge dilemmas

When comparing the standard switch and standard footbridge dilemmas in all clusters for the trolley and the speedboat tasks we found good enough evidence in every case for the support of the alternative hypothesis. The summary results of each comparison separately can be found in Table X.

Oxford utilitarianism Scale

```
# Prepare data for plotting
plot_data <-
  trolley_proc %>%
  select(survey_name, lab, ResponseId, contains("oxford_")) %>%
  pivot_longer(
    contains("oxford_"),
    names_to = "item_no",
    values_to = "rate",
    names_prefix = "oxford_utilitarian_") %>%
  mutate(item_no = as.integer(item_no),
         subscale = case_when(item_no == 1L ~ "Impartial Beneficence",
                              item_no == 2L ~ "Instrumental Harm",
                              item_no == 3L ~ "Impartial Beneficence",
                              item_no == 4L ~ "Instrumental Harm",
                              item_no == 5L ~ "Impartial Beneficence",
                              item_no == 6L ~ "Instrumental Harm",
                              item_no == 7L ~ "Impartial Beneficence",
                              item_no == 8L ~ "Instrumental Harm",
                              item_no == 9L ~ "Impartial Beneficence",
                              TRUE ~ NA_character_),
         survey_name = str_remove(survey_name, "PSA006_")) %>%
  drop_na(rate)

# Create plot
plot_data %>%
  ggplot() +
  aes(
    x = subscale,
    y = rate,
    width = 0.7) +
  geom_boxplot() +
  facet_wrap(facets = vars(survey_name)) +
  labs(
    y = "Rating",
```

```

x = "Subscale") +
scale_y_continuous(
  breaks = 1:7,
  labels = as.character(1:7)) +
theme_bw() +
theme(
  panel.grid.major = element_blank(),
  panel.grid.minor = element_blank(),
  plot.title = element_text(hjust = 0.5, vjust = 0.5, face = "bold", size = 17),
  axis.title.x = element_blank(),
  axis.title.y = element_text(face = "bold", size = 10),
  axis.text.x = element_text(face = "bold", size = 7, colour = "Black"),
  axis.text.y = element_text(face = "bold", size = 12, colour = "Black"),
  legend.title = element_blank(),
  legend.text = element_text(size = 9))

```

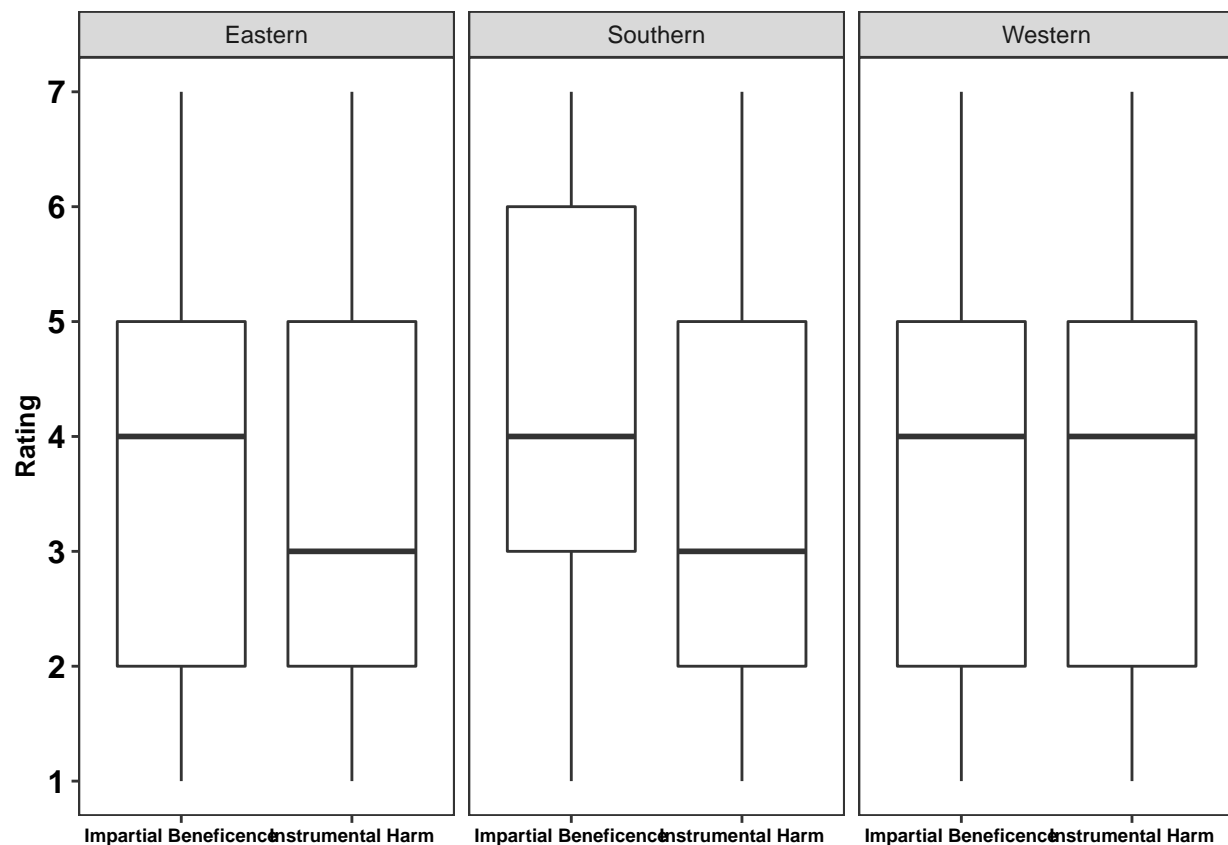


Table 1

```

### Some country codes are not standard iso codes; replace
custom_countries <-
  c("LEB" = "LBN",
    "BUL" = "BGR",
    "SPA" = "ESP",
    "SWT" = "CHE",

```

Eastern

China, India, Iran, Japan, Lebanon, Malaysia, North Macedonia, Pakistan, Thailand, United Arab Emirates, United Kingdom

```
"PSA" = "USA")

# Transform data for the table
## Table 1 represents the countries who collected data for the study
## in the three investigated clusters
table_data <-
  trolley_raw %>%
  select(survey_name, lab) %>%
  drop_na(lab) %>%
  mutate(country3 = str_extract(lab, "[A-Z]+") %>%
    recode(., !!!custom_countries),
    region = str_remove(survey_name, "PSA006_"),
    Country = countrycode(sourcevar = country3,
      origin = "iso3c",
      destination = "country.name"))

# GBR_001!
# Check the number of countries per region
table_data %>%
  distinct(region, Country) %>%
  count(region)

##      region  n
## 1 Eastern 11
## 2 Southern 12
## 3 Western 24

# Missing countries since the Stage 1 manuscript:
# Western: South Africa
# Southern: El Salvador
# Eastern: Indonesia

table_data <-
  table_data %>%
  distinct(region, Country) %>%
  group_by(region) %>%
  arrange(Country) %>%
  summarise(countries = str_c(Country, collapse = ", ")) %>%
  pivot_wider(names_from = region, values_from = countries)

## `summarise()` ungrouping output (override with `.groups` argument)

# Create APA formatted table
papaja::apa_table(
  table_data,
  caption = "The Cultural Classification of Countries of Participating Labs Following Awad et al.",
  escape = TRUE
)
```