ACrep\_Results.Rmd

Hannah Moshontz, running Erin Buchanan’s Analysis Script … final updates by Erin Buchanan

Last Updated 2022-06-08

# Run the analysis script

## Information

Multilab close replication of: Experiment 1 from Turri, J., Buckwalter, W., & Blouw, P. (2015). Knowledge and luck. *Psychonomic Bulletin and Review*, 22, 378-390.

[Data and registered protocols:](https://osf.io/n5b3w/)

[Codebook (raw data)](https://docs.google.com/spreadsheets/d/1KjXqgfVgguHeDXVtlHHhJ9zsRGVDfPVPH4tbh75P46U/edit#gid=903093128)

[Preprint](https://psyarxiv.com/zeux9/)

## Libraries

## Import the Data

The full\_long dataset includes all participants in long format - wherein each trial of their study is on one row of the dataset. Our uploaded data also includes full.Rds which is the same data in wide format - one column for each variable in the dataset and one row per participant.

## Open Ended Response Coding

Import the open-ended responses and prepare them for merging with the final dataset. This section:

* Imports all the files
* Includes a coder id
* Corrects all the issues with typed codings
* Converts each item to coder counts so you can convert from wide to long
* Converts from wide to long
* Converts into scoring:
  + No or NA gets 0 points
  + Maybe gets 1 point
  + Yes/test gets 2 points
  + Participants are marked as “excluded” if they get 4 total points across three coders
* Participants are also marked as “nonsense” if they did not write a legible answer or simply typed gibberish. These data points are not excluded but marked if you would like to repeat the analysis and exclude them.

### Finding missing coders:

## Fix Data Coding / Examine Covariates

### Deal with lab coding issues:

* Import lab sheet
* Delete cases that should not be used

### Deal with compensation issues:

* Deal with compensation for labs with no information because they didn’t see that question

### Fix Lab ID to Consistent “Data Collection Location ID”

### Examine Coding for Covariates

* Examine gender labels
* Recode education to number and examine … fix issues with two labs not using the same education values.

## Study Exclusions

In this section, we will mark each exclusion criteria to be able to denote how many exclusions each participant may have had.

### Majority Age

1. if the participant is not the majority age of their country or older (unless parent/guardian waiver provided)

In this section, we will exclude all participants who are under 18 or are not the majority age of their country or older.

<https://en.wikipedia.org/wiki/Age_of_majority>

### Previous Knowledge/Participation of Study

1. if the participant has taken part in a previous version of this study or in another contributors’ replication of the same study

### Comprehension Questions

1. if the participant fails to answer comprehension questions correctly

### Knowledge of Experiment/Hypothesis

1. if the participant correctly and explicitly articulate knowledge of the specific hypotheses or specific conditions of this study when answering the funneled debriefing questions.

### Language Proficiency

We will also exclude participants who self-report their understanding of the tested language as “not well” or “not well at all”. We based this exclusion criteria on a recent study that found that non-native English speakers who self-report as “very well” and “well” tend to score in the “intermediate” and “basic” categories on an English proficiency test respectively, while those who self-report as “not well” and “not at all” tend to score in the “below basic” category (Vickstrom, Shin, Collazo, & Bauman, 2015). All excluded data will be included in the data files on the overall OSF page, along with the particular reason for why they were excluded.

### Total Exclusions

In this section, we provide the total (data point) exclusions for each particular screening item plus a table with all exclusions forever.

## Examine the Distributions

## Descriptives and Demographics

## Analysis Overview

* Analysis:
  + Multilevel modeling using appropriate linking function
* Dependent variable(s):
  + Knowledge
  + Luck
  + Reasonable
* Independent variable:
  + Condition
* Covariates:
  + Test setting: compensated versus uncompensated
  + Age
  + Gender (men/women/other)
  + Years of education
* Random intercepts:
  + participant / vignette / lab / maybe geographic region
* Special population:
  + Turk data on similar analysis

## Data Screening

In this section, we will screen the continuous VAS variable, as the BIN variable does not require the same screening because of the different assumptions.

### Knowledge

### Reasonableness

### Luck

### Update DV

In this instance, we do not see any form of linearity or homoscedasticity - the data is clearly skewed or bimodal. Therefore, we decided to split the VAS scale into 40 and below and 60 and above, then combine with the truly binary data, and include this variable to understand if there were differences when the VAS scale did not produce continuous data.

Please note that the VAS is scored in the opposite direction of the BIN scales, therefore, you will see that the data split corrects this directionality difference.

We recoded knowledge and reasonableness for easier processing in the paper.

### Gender Other

Because of the small other category, we will simply compare male to female.

## Data Checks

### Missing Data Information

### Check for Reverse Coding

## Final Analysis Knowledge

### Model 1

Intercept only model to help determine if the addition of random intercepts are useful.

### Model 2

Vignette random intercept model:

* Remember that 0 is knows and 1 is believes
* Model 1 AIC: 18881.0942715
* Model 2 AIC: 17834.749218
* Interpretation: the inclusion of the vignette random variable is useful
* Coefficients: interpretation overall of random intercept, which effectively tells us regardless of condition, which way the participants lean toward answering. See the table printed out to help understand these values.
  + Darrel: fairly even split for knows/believes with a slight favor to knows (because it’s negative goes toward the 0 group)
  + Emma: leans more heavily toward the believes category (almost 4 to 1 odds)
  + Gerald: leans slightly toward the believes category
  + The results here indicate what we were seeing with the “continuous” data as well - that there’s a big difference in the way they respond to vignettes
* This vignette accounts for 8 to 10 percent of the random variance.

### Model 3

Vignette nested in participants random intercept:

* Model 2 AIC: 17834.749218
* Model 3 AIC: 17836.7493889
* Interpretation: the inclusion of participant ID isn’t really useful, as the AIC is higher for this model. The variance between participants is very small after accounting for vignettes.
* This model accounts for approximately the same random variance.

### Model 4

Vignette nested in participants which is then nested in labs:

* Model 2 AIC: 17834.749218
* Model 3 AIC: 17836.7493889
* Model 4 AIC: 17838.7492181
* Interpretation: the inclusion of lab ID isn’t really useful, as the AIC is higher for this model. The variance parameter for lab is also very small.
* This model accounts for approximately the same random variance.
* We will continue to use this cross classified structure to ensure we account for correlated error and to follow the pre-print.

### Model 5

Addition of the covariates added to the vignette random intercept model.

* Model 4 AIC: 17838.7492181
* Model 5 AIC: 17554.3090714
* Interpretation: this model is better than the vignette random intercept only model.
* Age is a negative predictor of knows/believes, such that as age increases, we are more likely to pick knows.
* Education is a positive predictor, so that as education increases we are more likely to pick believes.
* The fixed effects are very small ~ .001.

### Model 6

The previous model of covariates plus the variable of focus: condition.

* Model 5 AIC: 17554.3090714
* Model 6 AIC: 15871.9909757
* Interpretation: this model is better than the model with the covariates.
* Both knowledge and ignorance are different than the Gettier condition.
* In Gettier, participants are slightly more likely to say believes.
* In Ignorance, participants are *way* more likely to say believes (so that’s why this condition is different from Gettier).
* In Knowledge, participants are more like to say knows (also why this is different from Gettier).
* This effect accounts for 12-15% of the fixed effect variance.

### Model 6A

Does the vignette interact with the condition?

* Model 6 AIC: 15871.9909757
* Model 6A AIC: 15807.6929319
* Interpretation: This model is better than a model without the interaction.
* There is an interaction between vignette and condition for choice.
* In looking at the graph, we can see this is driven by the Emma vignette which shows a larger believes box for all conditions.
* This interaction adds a good portion of variance to the fixed effects 7-9%.

To analyze this interaction, we will use chi-square analysis and proportion tests (making sure it’s on the same data as the analysis above).

* The proportions are the likelihood of Knows for each condition
* Knows more likely in Gettier than Ignorance
* Knows more likely in Knowledge than Gettier
* Knows more likely in Gettier than Ignorance
* Knows more likely in Knowledge than Gettier
* Knows more likely in Gettier than Ignorance
* Knows more likely in Knowledge than Gettier
* While all three of these show the same results, we see that their strengths are different for each one using Cramer’s V.

### Model 6B

Data was collected from mechanical turk participants. Are there differences in this data and the complete data on the condition variable?

* Model 6 AIC: 15871.9909757
* Model 6A AIC: 15850.1573048
* Interpretation: This model is better than a model without the interaction.
* However, the interaction with the turk variable does not produce significant effects.
* The model is likely “better” because of the added variable and main effect of turk (likely drive by sample size differences in the two samples of turk versus not for knows versus believes).
* Therefore, we do not find evidence for differences in the turk sample for condition effects.

### Model 6C

Data was binned because of the distributions of the visual analogue scale - does the data source make a difference?

* Model 6 AIC: 15871.9909757
* Model 6C AIC: 15876.5654988
* Interpretation: This model is not better than a model without the condition interaction.
* There is no evidence for an interaction between condition and the data source, indicating that the binning process did not give us evidence of different results.

## Final Analysis Reasonableness

### Model 1

Intercept only model to help determine if the addition of random intercepts are useful.

### Model 2

Vignette random intercept model:

* Remember that 0 is reasonable, 1 is unreasonable
* Model 1 AIC: 7343.3496244
* Model 2 AIC: 7286.5547457
* Interpretation: the inclusion of the vignette random variable is useful
* Coefficients: interpretation overall of random intercept, which effectively tells us regardless of condition, which way the participants lean toward answering. See the table printed out to help understand these values.
  + We see that they all lean towards reasonable at slightly different strengths.
  + The intercept accounts for only a small amount of random variance, < .01 to .02.

### Model 3

Vignette nested in participants random intercept:

* Model 2 AIC: 7286.5547457
* Model 3 AIC: 7288.555639
* Interpretation: the inclusion of participant ID isn’t really useful, as the AIC is higher for this model. The variance between participants is very small after accounting for vignettes.
* This model accounts for approximately the same random variance.

### Model 4

Vignette nested in participants which is then nested in labs:

* Model 2 AIC: 7286.5547457
* Model 3 AIC: 7288.555639
* Model 4 AIC: 7290.5546228
* Interpretation: the inclusion of lab ID isn’t really useful, as the AIC is higher for this model. The variance parameter for lab is also very small.
* This model accounts for approximately the same random variance.
* We will continue to use this cross classified structure to ensure we account for correlated error and to follow the pre-print.

### Model 5

Addition of the covariates added to the vignette random intercept model.

* Model 4 AIC: 7290.5546228
* Model 5 AIC: 7144.101993
* Interpretation: this model is better than the vignette random intercept only model.
* Coefficients:
  + The compensated participants were more likely to say reasonable (they are also a much larger group).
  + Gender - females are more likely to pick reasonable (13:1 odds) versus men (11:1 odds)
  + Education is a negative coefficient indicating that the less education a person listed, the more likely they were to pick reasonable.

### Model 6

The previous model of covariates plus the variable of focus: condition.

* Model 5 AIC: 7144.101993
* Model 6 AIC: 7047.128759
* Interpretation: this model is better than the model with the covariates.
* Both knowledge and ignorance are different than the Gettier condition.
* In Ignorance, participants are likely to say reasonable … but in Gettier, they are more likely to say reasonable.
* In Knowledge, participants are more like to say reasonable than the Gettier condition.
* This effect accounts for 1-4% of the fixed effect variance.

### Model 6A

Does the vignette interact with the condition?

* Model 6 AIC: 7047.128759
* Model 6A AIC: 7025.8049336
* Interpretation: This model is better than a model without the interaction.
* There is an interaction between vignette and condition for choice.
* In looking at the graph, we can see this is driven by the Emma vignette which shows a larger unreasonable box for all conditions.
* This interaction adds a portion of variance to the fixed effects 1-5%.

To analyze this interaction, we will use chi-square analysis and proportion tests (making sure it’s on the same data as the analysis above).

* The proportions are the likelihood of Reasonable for each condition
* Reasonable more likely in Gettier than Ignorance
* No differences for reasonable in the Gettier and Knowledge Condition
* No differences in Gettier and Ignorance
* Reasonable more likely in the Knowledge Condition than Gettier
* Reasonable more likely in the Gettier than the Ignorance
* No differences in the Geitter and the Knowledge conditions

### Model 6B

Data was collected from mechanical turk participants. Are there differences in this data and the complete data on the condition variable?

* Model 6 AIC: 7047.128759
* Model 6A AIC: 7017.3668549
* Interpretation: This model is better than a model without the interaction.
* However, the interaction with the turk variable does not produce significant effects.
* The model is likely “better” because of the added variable and main effect of turk (likely drive by sample size differences in the two samples of turk versus not for knows versus believes).
* Therefore, we do not find evidence for differences in the turk sample for condition effects.

### Model 6C

Data was binned because of the distributions of the visual analogue scale - does the data source make a difference?

* Model 6 AIC: 7047.128759
* Model 6C AIC: 7041.2887035
* Interpretation: This model is better than a model without the source interaction.
* There is no evidence for an interaction between condition and the data source, indicating that the binning process did not give us evidence of different results.

## Final Analysis Luck

### Model 1

Intercept only model to help determine if the addition of random intercepts are useful.

### Model 2

Vignette random intercept model:

* Remember that 0 is luck and 1 is ability
* Model 1 AIC: 17776.6718312
* Model 2 AIC: 16771.3040331
* Interpretation: the inclusion of the vignette random variable is useful
* Coefficients: interpretation overall of random intercept, which effectively tells us regardless of condition, which way the participants lean toward answering. See the table printed out to help understand these values.
  + Darrel: People are more likely to pick ability
  + Emma: People are more likely to pick luck
  + Gerald: People are more even, but lean toward luck
  + The results here indicate what we were seeing with the “continuous” data as well - that there’s a big difference in the way they respond to vignettes
* This vignette accounts for 8 to 10 percent of the random variance.

### Model 3

Vignette nested in participants random intercept:

* Model 2 AIC: 16771.3040331
* Model 3 AIC: 16773.3040807
* Interpretation: the inclusion of participant ID isn’t really useful, as the AIC is higher for this model. The variance between participants is very small after accounting for vignettes.
* This model accounts for approximately the same random variance.

### Model 4

Vignette nested in participants which is then nested in labs:

* Model 2 AIC: 16771.3040331
* Model 3 AIC: 16773.3040807
* Model 4 AIC: 16775.3043138
* Interpretation: the inclusion of lab ID isn’t really useful, as the AIC is higher for this model. The variance parameter for lab is also very small.
* This model accounts for approximately the same random variance.
* We will continue to use this cross classified structure to ensure we account for correlated error and to follow the pre-print.

### Model 5

Addition of the covariates added to the vignette random intercept model.

* Model 4 AIC: 16775.3043138
* Model 5 AIC: 16489.5956655
* Interpretation: this model is better than the vignette random intercept only model.
* Coeffients:
  + Compensation: when people are compensated they are more likely to pick luck.
  + Age: older individuals are more likely to pick ability.
  + Education: this effect indicates that people who select lower education levels are more likely to select luck versus ability. The variance accounted for is small.

### Model 6

The previous model of covariates plus the variable of focus: condition.

* Model 5 AIC: 16489.5956655
* Model 6 AIC: 15896.1678952
* Interpretation: this model is better than the model with the covariates.
* Both knowledge and ignorance are different than the Gettier condition.
* In Gettier, people are more likely to chose luck.
* In Ignorance, people are more likely to chose ability.
* In Knowledge, people are more likely to chose ability.
* This effect accounts for 5-6% of the fixed effect variance.

### Model 6A

Does the vignette interact with the condition?

* Model 6 AIC: 15896.1678952
* Model 6A AIC: 15458.374183
* Interpretation: This model is better than a model without the interaction.
* There is an interaction between vignette and condition for choice.
* In looking at the graph, we can see that each vignette produces a different pattern of results.
* This interaction adds a good portion of variance to the fixed effects 14-17%.

To analyze this interaction, we will use chi-square analysis and proportion tests (making sure it’s on the same data as the analysis above).

* The proportions are the likelihood of Luck for each condition
* No differences in Gettier and Ignorance
* More likely to pick luck in Gettier than Knowledge
* More likely to pick luck in Gettier than Ignorance
* More likely to pick luck in Gettier than Knowledge
* More likely to pick luck in Gettier than Ignorance
* More likely to pick luck in Gettier than Knowledge

### Model 6B

Data was collected from mechanical turk participants. Are there differences in this data and the complete data on the condition variable?

* Model 6 AIC: 15896.1678952
* Model 6A AIC: 15872.7277629
* Interpretation: This model is better than a model without the interaction.
* However, the interaction with the turk variable does not produce significant effects.
* The model is likely “better” because of the added variable and main effect of turk (likely drive by sample size differences in the two samples of turk versus not for knows versus believes).
* Therefore, we do not find evidence for differences in the turk sample for condition effects.

### Model 6C

Data was binned because of the distributions of the visual analogue scale - does the data source make a difference?

* Model 6 AIC: 15896.1678952
* Model 6C AIC: 15862.0881337
* Interpretation: This model is better than a model without the condition interaction.
* We do see evidence for some interaction with the source of the data.
* Gettier more likely to say Luck than Ignorance
* Gettier more likely to say Luck than Knowledge
* Gettier more likely to say Luck than Ignorance
* Gettier more likely to say Luck than Knowledge
* Interaction is likely driven by the differences in sample size and effect size for these effects (slightly stronger effects in the Binary condition but these overlap in their CI).

## Exploratory Analyses

### Condition by Luck

* Model 6 AIC: 15871.9909757
* Model 6E AIC: 13363.9778914
* Interpretation: This model is better than the original model without the interaction.
* There is an interaction between condition, luck choice to predict knowledge.
* When we pick luck:
  + Gettier more likely to pick knows than Ignorance
  + Knowledge more likely to pick knows than Gettier
* When we pick ability:
  + Gettier more likely to pick knows than Ignorance (by a lot)
  + Knowledge more likely to pick knows than Gettier (by a little)

### Alternative Knowledge Question

* Remember, 0 = knows, 1 = only feels
* Model 6 AIC: 15871.9909757
* Model 6F AIC: 16332.6777166
* Interpretation: this model is “better” than the original, in that the AIC is lower, indicating lower error in predicting the outcome.
* In the Gettier condition, people are more likely to choose only feels
* In the Ignorance condition, people are more likely to choose only feels but way more than the Gettier condition (why this predictor is significant)
* In the Knowledge condition, they are fairly equal at choosing knows versus feels, but which is this one is different than Gettier
* These results appear to be the same pattern as the original question

### Order Effects

* Model 6 AIC: 15871.9909757
* Model 6G AIC: 15840.958235
* Interpretation: this model is “better” than the original, in that the AIC is lower, indicating lower error in predicting the outcome. This result isn’t too surprising because we’ve added two new predictors.
* There do appear to be order effects for both predictors.
* However, these plots indicate that there’s likely not a lot of effect for order - only slightly differences in knows versus believes selection.
  + Model 6 AIC: 15840.958235
  + Model 6F AIC: 15858.0283606
  + Interpretation: This model is not better than a model with just the order effects.
  + There is no interaction between condition and order effects.

## Darrel Analysis

## Lab Information

# Set up

#knitr::opts\_chunk$set(cache = TRUE)  
library(janitor)

##   
## Attaching package: 'janitor'

## The following objects are masked from 'package:stats':  
##   
## chisq.test, fisher.test

library(stringr)  
library(kableExtra)

##   
## Attaching package: 'kableExtra'

## The following objects are masked from 'package:flextable':  
##   
## as\_image, footnote

## The following object is masked from 'package:dplyr':  
##   
## group\_rows

library(countrycode)  
library(ISOcodes)  
#library(tidylog)  
library(naniar) # to evaluate missing data  
library(broom.mixed) # for nicely presented results  
library(cowplot)  
library(RColorBrewer)  
library(flextable)   
library(dplyr)

#a function to make names in the tidy model objects less confusing  
#used to prep results tables  
clean\_variable\_names <- function(tidy\_model\_obj){  
   
 tidy\_model\_obj %>%   
 mutate(group = str\_replace(group, "person\_code.\*", "site"),  
 group = str\_replace(group, "id:vignette","participant")) %>%   
 mutate(term = str\_replace(term, "compYes", "compensation"),  
 term = str\_replace(term, "gender2male", "gender"))  
   
}

**Deviations from Provisionally Accepted Protocol**

The protocol for this study was accepted as a Stage 1 Registered Replication Report (<https://psyarxiv.com/zeux9/>; see Appendix X). In this section, we describe the method as implemented and deviations from the protocol, including minor adjustments to language, corrections of factual inaccuracies, and methodological alterations. The primary deviations from the approved protocol, albeit minor, consisted of changes to study procedure and the analysis plan, due to error and adaptations required for valid statistical inference. As detailed below, we made changes to the methodology in terms of how surveys were programmed and implemented, how we operationalized luck attribution, how we measured race/ethnicity, and how we evaluated statistical analyses conducted by student-led teams. A number of aspects were not sufficiently described in the original protocol; we therefore clarified the analysis plan in terms of exclusion criteria, dropping of covariates that indicated whether the study was conducted individually versus in a group setting and in-person vs. online, as well as data assumption checking procedures.

**Project Teams**

# number of unique person\_code by osf ids (some people had multiple projects) that are not NA  
lab\_sheet$unique\_project <- paste0(lab\_sheet$person\_code, lab\_sheet$link\_code)  
  
n\_labs\_applied <- length(unique(lab\_sheet$unique\_project[lab\_sheet$unique\_project != "NANA"]))  
  
# number of unique link\_person combination that have protocol approved  
n\_labs\_rejected <- n\_labs\_applied - length(unique(na.omit(lab\_sheet$unique\_project[lab\_sheet$protocol\_approved == "Yes"  
])))  
  
# number of unique person\_codes who were approved but didn't collect data  
n\_labs\_failed <- length(unique(na.omit(lab\_sheet$unique\_project[lab\_sheet$data\_collected == "No"])))  
  
# number of labs who "finished" crep  
n\_labs\_finished\_crep <- length(unique(na.omit(lab\_sheet$unique\_project[lab\_sheet$completed\_crep\_requirements == "Yes"])))  
  
#check assumption that each site had one data collection source...  
#error out if any site contributed data from both Qualtrics and SSS  
if (full\_long %>%   
 group\_by(person\_code, source) %>% count() %>%   
 group\_by(person\_code) %>% count() %>%   
 filter(n > 1) %>%   
 nrow() %>% .[1] != 0) STOP  
  
#save survey platform info  
platform <- full\_long %>%   
 group\_by(person\_code) %>%   
 summarize(survey\_platform = first(source)) %>%  
 ungroup() %>%   
 glimpse()

## Rows: 38  
## Columns: 2  
## $ person\_code <chr> "04aec6c128a6d0c7b156124a782a2e17", "053921f0049652387…  
## $ survey\_platform <chr> "SoSciSurvey", "SoSciSurvey", "SoSciSurvey", "SoSciSur…

#check assumption that each site had one compensation type  
multiple\_compensation <- lab\_sheet %>%   
 group\_by(person\_code, compensation) %>% count() %>%   
 group\_by(person\_code) %>% count() %>%   
 filter(n > 1) %>% pull(person\_code)  
  
# Erin remade this table  
use\_data <- subset(lab\_sheet, use == "Yes")  
full\_sample\_n <- as.data.frame(table(full\_long$lab\_id))  
colnames(full\_sample\_n) <- c("lab\_code", "full\_sample\_size")  
final\_sample\_n <- as.data.frame(table(final\_long$lab\_id))  
colnames(final\_sample\_n) <- c("lab\_code", "final\_sample\_size")  
use\_data <- merge(use\_data, full\_sample\_n, all.x = T, by = "lab\_code")  
use\_data <- merge(use\_data, final\_sample\_n, all.x = T, by = "lab\_code")  
use\_data$full\_sample\_size <- use\_data$full\_sample\_size / 3  
use\_data$final\_sample\_size <- use\_data$final\_sample\_size / 3

Each student-led project team prepared a study protocol for approval by a CREP reviewer to ensure quality control. Teams could not contribute to data collection until their protocol was approved. For more information about this process and detailed descriptions of logistical considerations, see Appendix A[[1]](#footnote-90). In total, 65 student-led teams (i.e., unique teams with OSF pages) signed up to collect data for this project. At one data collection site, multiple student-led project teams completed the project (see Table @ref(tab:table1)). 51 student-lead project teams were approved to begin data collection using CREP procedure guidelines. Only 47 of these teams contributed to the full dataset, which represents 38 data collection sites. Teams not included in the full dataset either did not collect any data (e.g., due to campus closures during the COVID-19 pandemic) or because data were collected in a way that rendered them unusable for analyses (e.g., unable to match to full data, vignettes were not properly randomized). After applying exclusions described below, the final dataset includes 45 student-led project teams across 37 data collection sites. In total, 22 teams completed all CREP requirements including site level analysis and other pedagogical tasks.

Table 1. *Characteristics of project teams*

flextable(THAT\_TABLE)

| person\_code | Language | GP\_Region | Site | Teams | Sample | In\_Person | In\_Group | Compensation | Site\_Analysis | Full\_N | Final\_N |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 04aec6c128a6d0c7b156124a782a2e17 | English | USA |  | 1 | Undergraduates | Yes | Unclear | Course Credit |  | 387 | 221 |
| 053921f004965238755a0b08414ad61f | English | CAN |  | 1 | Undergraduates | No | No | Course Credit |  | 551 | 258 |
| 10fe64c77bdcdd1f63c4f6f6767bd834 | Chinese | TWN |  | 1 | Community | No | No | Money |  | 452 | 89 |
| 1300dd2a365669443e58645538040b7e | English | GRC |  | 1 | Both | No | No | None |  | 98 | 52 |
| 20e15d6f0eb0520239a9f7921b1ee905 | Turkish | TUR |  | 1 | Undergraduates | No | No | Course Credit |  | 304 | 77 |
| 25c0099bdeb9900e64b3f1e17ea9cf32 | Polish | POL |  | 1 | Unclear | Unclear | Unclear | None |  | 179 | 72 |
| 2c70764140dc2267dc65e463842c75dd | Hungarian | HUN |  | 1 | Undergraduates | Both | Unclear | Course Credit |  | 658 | 449 |
| 3b4e9381f4aafb032e37c0d5fa732aba | English | SGP |  | 1 | Undergraduates | Yes | Yes | Course Credit |  | 78 | 52 |
| 3e1f2a5e94450c8371562c74671ace76 | English | NZL |  | 1 | Unclear | Yes | No | Lottery |  | 58 | 42 |
| 47cce7454402254dd1e56c7338d2ca3b | Portuguese | POR |  | 1 | Both | No | No | Course Credit for Students, Nothing for Community |  | 149 | 81 |
| 4f86a6d4b553a44559cde3ceba92b330 | French | CHE |  | 1 | Unclear | Unclear | Unclear | Course Credit, Money |  | 58 | 28 |
| 5f4f444af13a37ac3598f3f4f613f560 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 129 | 56 |
| 6109cf70459f0edc57ccb89cf26d66ef | German | AUT |  | 1 | Unclear | Unclear | Unclear | None |  | 159 | 76 |
| 63947e8c9e47bc0056f78f402f3898f1 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 103 | 36 |
| 64ba90ce7822b9c898df033ebb8c2c18 | German | AUT |  | 1 | Unclear | Unclear | Unclear | Unclear |  | 143 | 75 |
| 71289e5ce6b2abe28c0d25bdce7c14a0 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 162 | 93 |
| 7898bca2938d7b6d8a82093ddffb15dc | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 213 | 103 |
| 7f9c8bee305f2a759f2905cc5afecd6a | English | AUS |  | 1 | Undergraduates | Unclear | Unclear | Course Credit |  | 165 | 119 |
| 82d647ec9366dd067a092145d7d8ee7b | English\_English | GBR\_GBR |  | 1 | Undergraduates\_Undergraduates | No\_No | No\_No | Course Credit\_Course Credit |  | 340 | 132 |
| 8ceb72b59172af0d6a8d657e76ec4ce8 | English | AUS |  | 1 | Undergraduates | No | No | Course Credit |  | 215 | 122 |
| 8ee3ca908c989e3f6342994da802c3bd | Russian | RUS |  | 1 | Both | No | No | Course Credit |  | 233 | 99 |
| a5e2ba5c7fcf9a90c4bc1b5ed666f57b | Polish\_Polish | POL\_POL |  | 1 | Both\_Both | No\_No | No\_No | None\_None |  | 243 | 121 |
| adaccbbc254a881db6c19b3224887bcc | English | USA |  | 1 | Community | No | No | Money |  | 510 | 356 |
| b6bab6f7c2e61741dcdc1689c30fc5fb | English | USA |  | 1 | Undergraduates | Unclear | Unclear | Unclear |  | 8 |  |
| b975583369996d9136088dbe211abfcf | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 124 | 57 |
| bc3266a9a2dc0de236029a38f1e0d4b9 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 95 | 54 |
| bfe85dfa8a47470e2600b4bae1dfe4c1 | German\_German\_German\_German\_German\_German\_German | DEU\_DEU\_DEU\_DEU\_DEU\_DEU\_DEU |  | 7 | Undergraduates\_Undergraduates\_Undergraduates\_Undergraduates\_Undergraduates\_Undergraduates\_Undergraduates | Yes\_Yes\_Yes\_No\_No\_No\_No | No\_No\_No\_No\_No\_No\_No | Sweets\_Sweets\_Sweets\_Money\_None\_None\_Lottery |  | 1,102 | 588 |
| c22875ee12b4cdfb8843ead3b8e1ddf8 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 164 | 93 |
| c6e892a926333f0840a81af6fa74a6c6 | Slovak | SVK |  | 1 | Both | No | No | None |  | 229 | 105 |
| ccc20a00466f0c4077f924733e9a3789 | Norwegian | NOR |  | 1 | Both | No | No | Sweets |  | 147 | 76 |
| d4c4b2302987480f7ce2084927489e4e | English | USA |  | 1 | Undergraduates | Yes | Unclear | Course Credit |  | 64 | 43 |
| d5450db323eeab9cf7bd57ecc95d3454 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 112 | 56 |
| d547b0fe851b1f8786fce767e3d89f50 | English\_English | USA\_USA |  | 1 | Undergraduates\_Undergraduates | Unclear\_Unclear | Unclear\_Unclear | Course Credit\_Course Credit |  | 402 | 201 |
| de5a43985156a022e44511202ccd5449 | Romanian | ROU |  | 1 | Undergraduates | No | No | Course Credit |  | 661 | 371 |
| ec35cfeaedd64eccef16c265b2ba179f | German | DEU |  | 1 | Undergraduates | No | No | Course Credit |  | 332 | 184 |
| ecb1153cd73ee30290a753fb74358c19 | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 91 | 48 |
| f19f790831cbe3722ccd98fd63d3d902 | English | USA |  | 1 | Both | No | No | Course Credit for Students, Nothing for Community |  | 135 | 63 |
| fd5a38067cf64a3e10ae8e92f015be4e | English | USA |  | 1 | Undergraduates | No | No | Course Credit |  | 187 | 78 |

*Note.* Site level analyses were conducted independently and not included in the analyses presented here but can be found on their OSF pages.

\*Site collected data using Qualtrics instead of SoSciSurvey

**Participants**

#create dataset for demographics   
final\_demographics <- final\_long %>%   
 group\_by(id) %>%  
 slice(1) %>%   
 ungroup() %>%   
 select(id,   
 source,  
 contains("exclusion"),  
 contains("nonsense"),  
 purpose,  
 previous,  
 survey\_lang,  
 turk,  
 language,  
 lab\_id,   
 person\_code,  
 lab\_country,  
 country,  
 un\_region,  
 gender,   
 gender\_other,  
 starts\_with("ethn\_"),  
 education,  
 education\_level,  
 starts\_with("comp"),  
 birth\_country,  
 case,  
 age)  
  
full\_demographics <- full\_long %>%   
 group\_by(id) %>%  
 slice(1) %>%   
 ungroup() %>%   
 select(id,   
 source,  
 contains("exclusion"),  
 contains("nonsense"),  
 purpose,  
 previous,  
 survey\_lang,  
 turk,  
 language,  
 lab\_id,   
 person\_code,  
 lab\_country,  
 country,  
 un\_region,  
 gender,   
 gender\_other,  
 starts\_with("ethn\_"),  
 education,  
 education\_level,  
 starts\_with("comp"),  
 birth\_country,  
 case,  
 age)  
  
#error out if count doesn't match  
if (length(unique(final\_long$id)) != nrow(final\_demographics)) STOP  
  
sites\_fewer\_than\_50 <- final\_long %>%   
 group\_by(lab\_id, id) %>%   
 count() %>%   
 group\_by(lab\_id) %>%   
 count() %>%   
 filter(n < 50) %>%   
 nrow()

In the analysis sample (i.e., after the exclusions described below), participants were 4826 adults recruited to participate by student researchers at 37 data collection sites in various geopolitical contexts across geographical regiοns (i.e., Northern America, Eastern Europe, Western Europe, Northern Europe, Southern Europe, Australia and New Zealand, Western Asia, Southeastern Asia, Eastern Asia). See Table @ref(tab:table2) for sample sizes by region. Data collection took place between January 1, 2019 and June 1, 2021[[2]](#footnote-91). Data collection sites contributed a median of 81 participants to the analysis sample (min = 28, max = 588); 8 sites collected fewer than the target of 50 participants. On average, participants were young (Mage = 24.84, SD = 9.91, n = 4826) and had completed some college as measured by years of education (Meducation = 13.84, SD = 2.59, n = 4771)[[3]](#footnote-92). Most participants (70.37; n = 3396) identified as White[[4]](#footnote-93). Over half of participants identified as female (70.56%; n = 3405) and most other participants identified as male (29.01%; n = 1400; n neither = 21). The plurality of participants completed the survey in English (47.53%; n = 2294). Participation details, like compensation and the sampled population, varied by data collection site. See Table @ref(tab:table1) for summary.

Table 2. *Number and Percent of Participants in the Analysis Dataset by Geopolitical Region*

flextable(  
 final\_long %>%   
 group\_by(lab\_country) %>%   
 summarize(n = n()/3,   
 percent = format(n() / nrow(final\_long)\*100, digits = 2, nsmall = 2)) %>%   
 arrange(desc(n))  
)

| lab\_country | n | percent |
| --- | --- | --- |
| USA | 1,501 | 31.10 |
| DEU | 666 | 13.80 |
| HUN | 449 | 9.30 |
| ROU | 371 | 7.69 |
| CAN | 258 | 5.35 |
| AUS | 241 | 4.99 |
| POL | 193 | 4.00 |
|  | 163 | 3.38 |
| AUT | 151 | 3.13 |
| GBR | 132 | 2.74 |
| SVK | 105 | 2.18 |
| RUS | 99 | 2.05 |
| TWN | 89 | 1.84 |
| PRT | 81 | 1.68 |
| TUR | 77 | 1.60 |
| NOR | 76 | 1.57 |
| GRC | 52 | 1.08 |
| SGP | 52 | 1.08 |
| NZL | 42 | 0.87 |
| CHE | 28 | 0.58 |

*Note.* Geopolitical region refers to the location of the data collection site except for one team that collected data through Amazon Mechanical Turk (MTurk) in another geopolitical region. For all other data collection sites, participants were recruited from the geopolitical region of the site.

Table 3. *Number and Percent of Participants by Data Collection Context Variables*

tempDF <- data.frame(variable = c("Compensated for participation",   
 "Recruited through mTurk",   
 "Took the centralized survey"),   
 n = c(unname(table(final\_long$comp)["Yes"]/3),   
 unname(table(final\_long$turk)["TRUE"])/3,   
 unname(table(final\_long$source)["SoSciSurvey"]/3)),   
 percent = c(unname(table(final\_long$comp)["Yes"])/nrow(final\_long),   
 unname(table(final\_long$turk)["TRUE"])/nrow(final\_long),   
 unname(table(final\_long$source)["SoSciSurvey"]/nrow(final\_long)))  
)  
tempDF$percent <- format(tempDF$percent\*100, digits = 2, nsmall = 2)  
  
  
flextable(tempDF)

| variable | n | percent |
| --- | --- | --- |
| Compensated for participation | 3,533 | 73.21 |
| Recruited through mTurk | 356 | 7.38 |
| Took the centralized survey | 4,270 | 88.48 |

*Note.* Variables are not exclusive. Information about the compensation method was obtained by examining each student-led team’s IRB approval, confirming with the students or PIs at each site, and making inferences based on the data collection site specific surveys when neither source was available. Two data collection teams included in analyses used Qualtrics to distribute their surveys instead of the centralized survey programmed in SoSciSurvey.

*Exclusions*

Of the 9440 participants who completed the survey, data from 48.9 (*n* = 4614) were excluded from the analytic sample. Of this total, 2187 participants were flagged for exclusion based on multiple criteria. All listed exclusions were pre-registered with one exception (i.e., maximum age)[[5]](#footnote-94). Participants were excluded for the following reasons:

* Age: the participant did not provide an age, listed an age greater than or equal to 100, or was not the majority age of their country or older, operationalized as at least 18 in all areas except Taiwan, where the age of consent is 20 (*n* = 2118; missing = 2040).
* Prior participation: the participant had taken part in a previous version of this study or in another contributors’ replication of the same study (*n* = 238).
* Comprehension: the participant failed to answer all three of the vignette comprehension questions correctly (*n* = 4376; missing = 1490; e.g., did not correctly identify whether Darrel was looking at a squirrel or a prairie dog).
* Knowledge of hypothesis: the participant correctly and explicitly articulated knowledge of the specific hypotheses or specific conditions of this study when asked what they thought the study hypothesis was (*n* = 203).
* Language proficiency: the participant reported their understanding of the language the survey was presented in as “not well” or “not well at all” (*n* = 2093; missing = 2003; see Vickstrom, Shin, Collazo, & Bauman, 2015).

**Evaluating Prior Participation and Knowledge of Study**. Evaluating the hypothesis exclusion criterion required subjective judgements about open-ended responses. Each non-missing observation was evaluated by three raters who spoke the language of the provided response. These three raters did not translate responses, but instead directly evaluated responses with respect to the exclusion criteria. [⅔ Rule]. See Appendix X for the instructions given to raters and information about how ratings were combined. Ratings themselves are available at <http://osf.io/gs29c>. Responses identified by raters as test cases (e.g., “TEST”) were excluded (n study purpose = 222; n previously participated = 170)[[6]](#footnote-96). Responses that were not coherent were labeled, but not excluded (n study purpose = 5; n previously participated = 3).

*Power Analysis*

We conducted an *a priori* power analysis, using the *powerCurve* function in the *simr* package (Green & MacLeod, 2016) in *R* to estimate the sample size required to detect an effect of knowledge condition on participants’ knowledge attributions with 90% power at α = .05[[7]](#footnote-97). To estimate the effect size, we considered 1) the effects observed in our pilot-test data (difference between Gettier and knowledge, β = 0.32; difference between Gettier and ignorance, β = -0.44), 2) both the difference between the Gettier condition and knowledge condition (Cramér’s V = .509) and the small non-significant difference between the Gettier condition and ignorance condition (which we manually calculated using the reported statistics; Cramér’s V = .16) from Experiment 1 of Turri et al. (2015), and 3) the Gettier intuition effect size from Starman and Friedman (2012; Experiment 1; Cohen’s *d* = 0.4). To be conservative, we selected a standardized fixed effect within the multilevel model analysis described below of .1 for our power analyses.

The model tested included random intercepts for data collection site, vignette, and participants, such that vignettes were nested within participants who were nested within sites. We simulated data using a standardized fixed effect regression parameter of .1. In these simulations, the number of participants per site was allowed to vary, but the number of vignettes (3) and the number of collection sites (9) were held constant. Results suggested that at least 32 participants per data collection site (i.e., 288 total participants; 864 total observations) would be required to detect an fixed effect regression parameter of .1 90% of the time with an alpha of .05. Considering the potential for attrition (e.g., due to lack of comprehension) and effect size heterogeneity between data collection sites (Kenny & Judd, 2019), we set a target sample size of 50 participants per data collection site. Of the 38 data collection sites, 37 met this target prior to exclusions.

**Materials and Measures**

As described in the approved protocol, we planned to collect all data using a single SocSciSurvey survey programmed to accommodate lab-specific variations. However, 8 student-lead teams used Qualtrics surveys that student researchers programmed themselves. The majority of the data collected via Qualtrics was not included in the full data set due to logistical challenges (e.g., no access to raw survey data); only 3 of the teams included in the analysis dataset used Qualtrics surveys. 7 student teams whose data were included in the analysis dataset collected data from all participants in-person (from an on-campus computer), 1 student teams collected data from some participants in-person, and for 12 student teams, whether data collection took place in-person was not clearly documented. All participants completed an online consent form. Deviations that labs made from the materials and procedure described here are summarized in Table X. All materials used in this replication, including the details of the vignettes and related pretests, are available in the Appendix and at osf.io/n5b3w.

*Vignettes*

In addition to the “Squirrel/Darrel” vignette from Turri et al. (2015), two vignettes were selected on the basis of their similarity to the original vignette, their quality, and their prevalence in the literature: the “Fake Barn/Gerald” vignette (Colaca et al., 2014; altered to more closely match the “Squirrel/Darrel” vignette), and the “Diamond/Emma” vignette (Nagel, San Juan, et al., 2013). The vignettes as administered in this study are reported in full in Appendix X. The vignettes were pretested for how effectively they manipulated the target construct and how much participants comprehended them (see Appendix X). Four student-lead teams included a fourth vignette that was not described in the approved protocol (see Larkin & Andreychik, 2019).

For each vignette, participants were randomly assigned without replacement to one of three conditions: a Gettier-type case condition in which the vignette subject correctly identified the target, but not due to the reason they thought it to be true (i.e., the “Threat” condition for Turri et al., 2015); a knowledge control condition in which the subject correctly identified the target due to their knowledge (i.e., the “No Threat” condition for Turri et al.); and an ignorance control condition in which the protagonist incorrectly identified the target, but not due to their knowledge (i.e., the “No Detection” condition for Turri et al.).

*Dependent Measures*

After each vignette, two primary and two exploratory dependent variables were measured. In line with the approved protocol, the majority of student-led teams (33 teams) included only the default visual analog scale ranging from 0 to 100 for three of these variables. Three teams participated in an optional extension that randomly assigned participants to take the study using either entirely continuous scale measures or entirely binary choice measures for these variables[[8]](#footnote-98). Overall, for each of the three measures, 86.52 of responses used in analyses were originally measured on the continuous scale. Exact question text is reported in Appendix X.

**Knowledge Attribution**. Participants were asked whether the protagonist believes or knows the stated proposition.

**Reasonableness Judgments**. Participants were asked to rate the extent to which the protagonist’s belief was unreasonable or reasonable.

**Luck Attribution**. For this exploratory measure, participants were asked two questions relevant for evaluating their attributions of luck or ability. First, participants were asked whether or not the protagonist got the “right” or “wrong” answer (binary choice in all cases). Then, participants were asked whether the protagonist’s “right” or “wrong” answer was due to their ability/inability or their good luck/ bad luck on one of the two scales. If participants selected the incorrect answer to the first part of the question, they were subsequently excluded from the luck attribution analyses only - given that their response indicated that they did not comprehend whether or not the protagonist held the given true belief.

**Alternative Knowledge Attribution**. In addition, participants were asked an alternative knowledge probe in which participants chose whether the protagonist either knew what the target of identification was or felt like they knew what the target was but did not actually know (binary choice in all cases). For example, in the Squirrel/Darrel scenario, participants were asked to select, “In your view, which of the following sentences better describes Darrel’s situation?” Participants could then select one of two response options: “Darrel knows that the animal he saw is a red speckled ground squirrel.” OR “Darrel feels like he knows that the animal he saw is a red speckled ground squirrel, but he doesn’t actually know that it is.”

*Demographics*

Participants were asked to report their age, gender, geopolitical region (i.e., “What country do you currently live in?”; “What is your country of birth?”), the number of years they had attended school, and their race or ethnicity. Because of differences in how student-led teams measured these items, we matched item answers across different implementations of the survey (described below).

**Education Level**. All participants were asked a question about their education. Participants who completed the survey in SocSciSurvey were asked about the number of years they had been in school (truncated at 18). Participants who completed the survey in Qualtrics were asked about their educational attainment. Education (in years) was imputed for participants who reported their educational attainment in these two sites (*n* = 553)[[9]](#footnote-99). The years of education for these two sites was also truncated to match how this item was measured in SocSciSurvey, such that any value above 17 was imputed as 18.

**Compensation.** Participants were asked whether or not they were compensated for their participation (“Will you receive any kind of compensation or reward for taking part in this study?”), and indicated the type of compensation (e.g., the number of course credits, the amount of money). Some student-led teams opted not to include this question in their survey because all participants were compensated the same way. The method of compensation described in the protocol was imputed for those missing responses for the compensation question for these reasons. Among participants who were asked about their compensation, responses were sometimes missing or discrepant with the documented method of compensation. For student-led teams where fewer than 50% of participants in the final dataset agreed on a method of compensation, the method of compensation described in the data collection site’s approved IRB protocol was imputed for all participants (if a single method of compensation was described).

*Participation and Exclusion Measures*

Participants were asked to indicate the true correct answer for each vignette as a comprehension check that we used for exclusions; the original study used the same type of question for the same purpose and excluded 15 of 135 participants on this basis. We also asked participants to describe what they thought the hypothesis of the study was (used for exclusions, see Appendix D for criteria), to provide their impression of study materials (not used in any analyses), and whether they had participated in a similar study (used for exclusions). The original study did not contain these questions, but the researchers completely excluded Amazon Mechanical Turk workers within and across the four studies in their 2015 paper if they had already participated. We also asked participants to rate their proficiency for the survey language (used for exclusions if participants responded “not well” or “not well at all”; see Vickstrom et al., 2015, for criteria). The original paper asked participants whether they were native English speakers, but did not seem to exclude participants on this basis. Given that the tasks in the present study were highly dependent on language comprehension and proficiency, and that participants had a 12.5% chance (i.e., 1 in 8) of passing all three comprehension questions based on guesses, we decided an additional check of self-reported language proficiency would be helpful in excluding participants who did not or may not have understood the task completely.

Participants also completed a 12-question study experience questionnaire that was not used in analyses (see Appendix X).

**Procedure**

After providing informed consent, participants read and answered questions about three vignettes that described counterfeit object cases. Each participant read and responded to one vignette for each protagonist and one for each knowledge condition. The random assignment of vignette and knowledge condition was executed by random presentation features in SocSciSurvey and Qualtrics. For each vignette, participants evaluated the protagonist’s knowledge on the knowledge attribution, reasonableness judgment, luck attribution, and other measures. Next, participants answered questions related to data exclusion criteria, demographics, and their experience completing the study. Finally, participants were debriefed and compensated if applicable.

**Analytic Approach**

Analyses were conducted on combined raw data collected in SocSciSurvey and Qualtrics. In the original protocol, we planned to evaluate the quality of each student-led team’s data, including the raw data, analysis scripts, codebooks, cleaned data sets, and narrative summaries of results. We also planned that data would be included in analyses only if these products passed a quality check. The original protocol did not describe clear criteria that would be used to detect and correct errors. In order to conduct reproducible, transparent analyses, we chose not to exclude data from teams who failed to meet the target sample size or did not submit their work for final CREP review. Because of the strict data quality exclusions implemented at the level of participant, we were not concerned about team level variation in data quality. A summary of how the teams independently analyzed their data (i.e., the test used for the effect of condition on knowledge attribution) is reported in Table @ref(tab:table1).

*Multilevel Models*

Multilevel models were used to evaluate our hypotheses. The unit of analysis was the question response, and cross-classified random intercepts for the vignette, participant, and data collection site were to be included to account for the nesting of responses within these groups. Exact model specification can be found at <https://osf.io/8ut6e/>[[10]](#footnote-101).

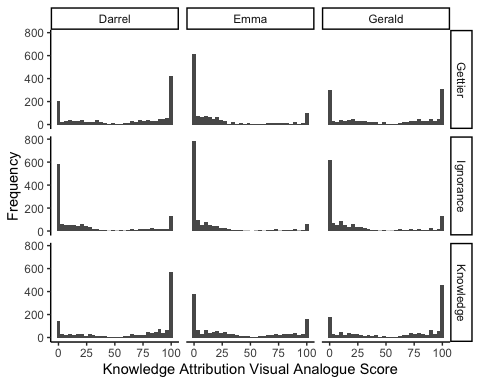
**Assumptions and Transformations**. While the approved protocol described testing assumptions before conducting analyses, it did not detail criteria that would be used for testing assumptions or approaches to handling model convergence issues. No convergence issues emerged during analyses. Here we describe the approach taken to test assumptions. Assumptions of and related to linearity are primarily relevant for the analysis of the continuously measured dependent variables. The continuous knowledge attribution variable was dramatically bimodal (see Figure 2).

*Figure 2*. Knowledge Attribution Visual Analogue Scores by Vignette and Condition

ggplot(final\_long, aes(know\_vas)) +   
 geom\_histogram() +   
 facet\_grid(cond~vignette) +   
 theme\_classic() +   
 ylab("Frequency") +   
 xlab("Knowledge Attribution Visual Analogue Score")

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 1952 rows containing non-finite values (stat\_bin).



Distributions of knowledge attribution by vignette and experimental condition were also dramatically bimodal. A linear model that predicted continuously measured knowledge attribution as a function of condition, with covariates of compensation, age, gender, and education, produced a bimodal residual distribution, indicating a violation of the residual normality assumption. Further, a plot of residuals by fitted values suggests that residuals vary as a function of predicted values, suggesting a violation of the homoscedasticity assumption. Similar patterns of results suggested that the assumptions of linear regression were also not met for continuously measured reasonableness (bimodal) or luck/ability variables (heavily skewed).

know\_vas\_trials\_dropped\_n <- final\_long %>%  
 filter(!is.na(know\_vas)) %>%   
 tabyl(know\_vas\_binned) %>%   
 filter(is.na(know\_vas\_binned)) %>%   
 pull(n)   
  
know\_vas\_trials\_dropped\_pct <- final\_long %>%  
 filter(!is.na(know\_vas)) %>%   
 tabyl(know\_vas\_binned) %>%   
 filter(is.na(know\_vas\_binned)) %>%   
 pull(percent) %>%   
 round(4) %>%   
 .[]\*100   
  
reason\_vas\_trials\_dropped\_n <- final\_long %>%  
 filter(!is.na(reason\_vas)) %>%   
 tabyl(reason\_vas\_binned) %>%   
 filter(is.na(reason\_vas\_binned)) %>%   
 pull(n)   
  
reason\_vas\_trials\_dropped\_pct <- final\_long %>%  
 filter(!is.na(reason\_vas)) %>%   
 tabyl(reason\_vas\_binned) %>%   
 filter(is.na(reason\_vas\_binned)) %>%   
 pull(percent) %>%   
 round(4) %>%   
 .[]\*100   
  
luck\_vas\_trials\_dropped\_n <- final\_luck %>%  
 filter(!is.na(luck\_vas)) %>%   
 tabyl(luck\_vas\_binned) %>%   
 filter(is.na(luck\_vas\_binned)) %>%   
 pull(n)   
  
luck\_vas\_trials\_dropped\_pct <- final\_luck %>%  
 filter(!is.na(luck\_vas)) %>%   
 tabyl(luck\_vas\_binned) %>%   
 filter(is.na(luck\_vas\_binned)) %>%   
 pull(percent) %>%   
 round(4) %>%   
 .[]\*100

Transforming continuous variables into discrete variables for analysis is not generally recommended (MacCallum et al., 2002; Maxwell & Delaney, 1993). For the present analyses, however, this approach was necessary due to the already bimodal distribution of the dependent variables. Thus, we split the continuously measured versions of the three dependent variables such that scores at and below 40 and scores at and above 60 were classified into discrete categories. Higher scores were coded as 0 to indicate knowledge, reasonableness, or luck, and lower scores were coded as 1 to indicate belief, unreasonableness, or ability. Using this approach resulted in minimal data loss. Of non-missing responses on each continuous measure, 359 (2.87) trials were dropped for the knowledge attribution variable, 279 (2.23) trials were dropped for the reasonableness attribution variable, and 683 (5.85) trials were dropped for the luck attribution variable. This approach allowed us to validly interpret model results and also test whether the method of measurement (continuous or binary) affected results.

**Model Steps**. A series of multilevel logistic regression models were fit predicting knowledge attributions, reasonableness judgements, and luck attributions. Because of the highly dichotomized responses to the continuous rating scale, we transformed those responses to a binary scale and analyzed them together with the responses originally measured on a binary scale. Each model was fit including all participants with no missing data on that model’s variables. After estimating a baseline intercept-only model (Model 1), we fit models with random intercepts for vignette (Model 2), person (Model 3), and data collection site (Model 4) added sequentially. In Model 5, participant age, compensation, gender, and education (in years) were added as fixed effects. These variables served as covariates and were included in our original analysis plan due to previous research that demonstrated their impact on knowledge attribution. Finally, the knowledge condition was added in Model 6. To see if the effect of condition varied by vignette, the interaction between vignette and condition was added as a fixed effect in Model 6A. Additional models were fit to test the moderating effects of participant source (Model 6B; mTurk vs. lab) and original measurement scale (Model 6C; binary vs. continuous). The conceptual models presented in 1-6B were pre-registered, maintaining independent and random effects variables in the updated analysis plan. As described, the dependent variable was dichotomized. Model 6C was added when the data screening indicated the VAS results were not continuous as expected. The technical details of the multilevel models (i.e., model order, interpretation) were updated from our pre-registered plan to ensure appropriate statistical inference.

# Results

To better test our research questions, we implemented analyses that differ from those we originally planned[[11]](#footnote-106). All deviations from the approved protocol are summarized in Table X. The results section as it appeared in the approved protocol is included in full in Appendix X with updated statistics where possible. While the results below indicate components of the random structure (i.e., intercepts of participant and site) do not add or improve the model, we included these facets to match the pre-registered plan and to maintain independence of observations (i.e., participant intercepts are arguably necessary for a repeated measures design). For each focal model, we report the model fit statistics and parameter estimates. Parameter estimates for logistic models can be interpreted in a similar fashion to linear regression models: negative values indicate that increasing the predictor decreases the likelihood of the dependent variable (e.g., the choice coded as one, therefore, increasing the likelihood of the choice coded as zero), and positive values indicate increasing the predictor correspondingly increases the likelihood of the dependent variable (e.g., the choice coded as one). When predictors are also categorical, increasing the predictor indicates a comparison between the predictor group coded as zero and the predictor group coded as one.

*Knowledge Attribution*

The goal of the present research was to provide a well-powered estimate of the magnitude and prevalence of Gettier intuitions (i.e., the difference in knowledge attribution between Gettier and knowledge conditions) across different scenarios and testing sites in a replication and extension of Turri et al., 2015. Models were fit in steps to determine whether participants attributed knowledge to the protagonist at different rates as a function of condition. Compared to the baseline Model 1 (AIC = 18881.09), the model including random intercepts for vignette (AIC = 17834.75) explained more variance. Pseudo-*R*2 values suggested that vignette accounts for 0.08 to 0.1 of the random variance in knowledge attributions. Inspection of frequencies indicated that participants attributed knowledge most frequently in response to the Gerald vignette (44.32) and least frequently in response to the Emma vignette (20.94).

k.model.2@frame %>%   
 group\_by(know\_vas\_combined, vignette) %>%   
 summarize(frequency = n()) %>%   
 arrange(desc(vignette)) %>%   
 pivot\_wider(id\_cols = c(know\_vas\_combined), names\_from = vignette, values\_from = frequency) %>%   
 kable(caption = "Knowledge Attributions by Vignette") %>%   
 kable\_styling()

## `summarise()` has grouped output by 'know\_vas\_combined'. You can override using  
## the `.groups` argument.

Knowledge Attributions by Vignette

know\_vas\_combined

Gerald

Emma

Darrel

0

2611

3716

2268

1

2077

991

2445

The model nesting vignette within participants (Model 3; AIC = 17836.75) explained similar amounts of variance (0.08 to 0.1) as Model 2. The addition of data collection site in Model 4 (AIC = 17838.7492181) likewise did not improve model fit (Pseudo-*R*2 = 0.08 - 0.1). The model including the covariates predicting knowledge attributions as fixed effects (Model 5; AIC = 17554.31) was more useful in explaining variance in knowledge attribution. Age negatively predicted the choice of knows versus believes, such that as age increases, participants were more likely to pick knows. Education was a positive predictor indicating that increases in reported education related to a higher likelihood of picking believes. However, these fixed effects were a small proportion of the variance, Pseudo-*R*2 < .001.

k.chance <- k.model.6@frame  
  
k.chance.i <- chisq.test(table(k.chance %>% filter(cond == "Ignorance") %>% select(know\_vas\_combined)))  
k.chance.g <- chisq.test(table(k.chance %>% filter(cond == "Gettier") %>% select(know\_vas\_combined)))  
k.chance.k <- chisq.test(table(k.chance %>% filter(cond == "Knowledge") %>% select(know\_vas\_combined)))

Model 6 served as the key replication test of Turri et al. (2015). The knowledge condition was added as a fixed effect (AIC = 15871.99). This model performed better than the previous model and revealed an effect of condition on knowledge attribution (Pseudo-*R*2 = 0.12 - 0.15). Participants were more likely to attribute knowledge to the protagonist in the knowledge condition vignette than to the protagonists in the ignorance and Gettier condition vignettes; further, the ignorance condition differed from the Gettier condition (see Table X). values suggest that condition accounted for of the fixed effect variance in knowledge attribution. Thus, we did not fully replicate the results of Turri et al. (2015, Experiment 1), who found no difference in knowledge attribution between the knowledge and Gettier conditions. Using the data only from this model, each condition was examined for difference from chance using . In the knowledge condition, the knows choice was more likely than chance, and the believes condition was more likely than chance in the Ignorance and Gettier condition, all *p*s < .001 (see Table X).

k.model.6@frame %>%   
 group\_by(know\_vas\_combined, cond) %>%   
 summarize(frequency = n()) %>%   
 arrange(desc(cond)) %>%   
 pivot\_wider(id\_cols = c(know\_vas\_combined), names\_from = cond, values\_from = frequency) %>%   
 kable(caption = "Knowledge Attributions by Condition") %>%   
 kable\_styling()

## `summarise()` has grouped output by 'know\_vas\_combined'. You can override using  
## the `.groups` argument.

Knowledge Attributions by Condition

know\_vas\_combined

Knowledge

Ignorance

Gettier

0

2005

3833

2638

1

2614

838

1967

*Does the effect of knowledge condition differ by vignette?*

k.dcond\_effect <- stats::chisq.test(data\_graph$know\_vas\_combined[data\_graph$vignette == "Darrel"],   
 data\_graph$cond[data\_graph$vignette == "Darrel"]) %>%   
 tidy()  
  
k.dcond\_GI <- prop.test(t(k.d\_table[2:1, 1:2])) %>%   
 tidy()  
  
k.dcond\_GI\_v <- v.chi.sq(x2 = prop.test(t(k.d\_table[2:1, 1:2]))$statistic,   
 n = sum(t(k.d\_table[2:1, 1:2])),  
 r = 2, c = 2)   
  
k.dcond\_GK <- prop.test(t(k.d\_table[2:1, c(1,3)])) %>%   
 tidy()  
  
k.dcond\_GK\_v <- v.chi.sq(x2 = prop.test(t(k.d\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(k.d\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

k.econd\_effect <- stats::chisq.test(data\_graph$know\_vas\_combined[data\_graph$vignette == "Emma"],   
 data\_graph$cond[data\_graph$vignette == "Emma"]) %>%   
 tidy()  
  
k.econd\_GI <- prop.test(t(k.e\_table[2:1, 1:2])) %>%   
 tidy()  
  
k.econd\_GI\_v <- v.chi.sq(x2 = prop.test(t(k.e\_table[2:1, 1:2]))$statistic,   
 n = sum(t(k.e\_table[2:1, 1:2])),  
 r = 2, c = 2)   
  
k.econd\_GK <- prop.test(t(k.e\_table[2:1, c(1,3)])) %>%   
 tidy()  
  
k.econd\_GK\_v <- v.chi.sq(x2 = prop.test(t(k.e\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(k.e\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

k.gcond\_effect <- stats::chisq.test(data\_graph$know\_vas\_combined[data\_graph$vignette == "Gerald"],   
 data\_graph$cond[data\_graph$vignette == "Gerald"]) %>%   
 tidy()  
  
k.gcond\_GI <- prop.test(t(k.g\_table[2:1, 1:2])) %>%   
 tidy()  
  
k.gcond\_GI\_v <- v.chi.sq(x2 = prop.test(t(k.g\_table[2:1, 1:2]))$statistic,   
 n = sum(t(k.g\_table[2:1, 1:2])),  
 r = 2, c = 2)   
  
k.gcond\_GK <- prop.test(t(k.g\_table[2:1, c(1,3)])) %>%   
 tidy()  
  
k.gcond\_GK\_v <- v.chi.sq(x2 = prop.test(t(k.g\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(k.g\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

To better understand whether the effect of condition varied as a function of the vignette’s content, Model 6A was estimated including an interaction between vignette and condition (AIC = 15807.69). This model explained more variance than Model 6. As shown in Figure X, the overall pattern of results was the same for every vignette; however, values suggest that the interaction between condition and vignette accounted for some of the variance in knowledge attributions. The size of the differences between conditions (and between vignettes) depended on the vignette-condition combinations (Pseudo-*R*2 = 0.2 to 0.24).

In responding to the Darrel vignette, participants attributed knowledge at different rates according to the vignette’s conditions, (2) = 781, p <.001. Participants were more likely to attribute knowledge when responding to the Gettier condition version (0.6) than in the ignorance condition version (0.23; Cramer’s = .37, 95% CI [.34, .41], (1) = 425.61, < .001). They were also more likely to attribute knowledge to Darrel when responding to the knowledge condition version (0.71 than in the Gettier condition version (0.6; Cramer’s = .12, 95% CI [.08, .15], (1) = 43.30, < .001).

The pattern of responding was similar for the Emma vignette; the likelihood that participants attributed knowledge to Emma differed according to the vignette’s condition, (2) = 291.42, p < .001. Participants were more likely to attribute knowledge when responding to the Gettier condition of the Emma vignette (0.18) than in the ignorance condition of the Emma vignette (0.1; Cramer’s = .11, 95% CI [.07, .14], (1) = 35.15, < .001). The likelihood of knowledge attribution was higher for the knowledge version of the vignette (0.35) than for the Gettier version (0.18; Cramer’s = .19, 95% CI [.16, .23], (1) = 112.59, < .001).

In response to the Gerald vignette, participant knowledge attributions similarly differed according to vignette condition, (2) = 607.03, p < .001. Participants were more likely to attribute knowledge in response to the Gettier condition version of the Gerald vignette (0.5) than to the ignorance condition version of the Gerald vignette (0.2; Cramer’s = .31, 95% CI [.28, .35], (1) = 304.67, < .001). In addition, they were more likely to attribute knowledge to Gerald in the knowledge condition version (0.63) than in the Gettier condition version (0.5; Cramer’s = .13, 95% CI [.09, .17], (1) = 50.27, < .001).

To interpret the condition by vignette interaction, we examined Cramer’s V for the analyses of each vignette. This approach revealed that the likelihood of knowledge attributions in the Gettier and ignorance conditions differed less for the Emma vignette than for the Darrel and Gerald vignettes. Additionally, the Gettier and knowledge conditions of the Darrel vignette produced a smaller difference in likelihood than that for those conditions of the other two vignettes. Thus, participants demonstrated Gettier intuitions in all three vignettes (i.e., participants were more likely to deny knowledge in the Gettier condition than in knowledge condition, a case justified true belief), but these Gettier intuitions were weakest in response to the Darrel vignette and strongest in response to the Emma vignette.

tidy(k.model.6A) %>%   
 clean\_variable\_names() %>%   
 kable(caption = "Interaction of Condition and Vignette on Knowledge Attributions") %>%   
 kable\_styling()

Interaction of Condition and Vignette on Knowledge Attributions

effect

group

term

estimate

std.error

statistic

p.value

fixed

NA

(Intercept)

0.5550196

0.1264092

4.3906588

0.0000113

fixed

NA

compensation

0.0167799

0.0448442

0.3741826

0.7082685

fixed

NA

age

0.0037784

0.0020039

1.8855192

0.0593598

fixed

NA

gender

-0.0876738

0.0427943

-2.0487271

0.0404888

fixed

NA

education

-0.0166002

0.0075607

-2.1955787

0.0281221

fixed

NA

condIgnorance

-1.6042917

0.0800233

-20.0478029

0.0000000

fixed

NA

condKnowledge

0.5031718

0.0761979

6.6034827

0.0000000

fixed

NA

vignetteEmma

-1.9304337

0.0844385

-22.8619962

0.0000000

fixed

NA

vignetteGerald

-0.3903202

0.0730638

-5.3421829

0.0000001

fixed

NA

condIgnorance:vignetteEmma

0.9774987

0.1327914

7.3611595

0.0000000

fixed

NA

condKnowledge:vignetteEmma

0.3950256

0.1145277

3.4491693

0.0005623

fixed

NA

condIgnorance:vignetteGerald

0.2231922

0.1137761

1.9616786

0.0497999

fixed

NA

condKnowledge:vignetteGerald

0.0217803

0.1061682

0.2051493

0.8374555

ran\_pars

site

sd\_\_(Intercept)

0.0000000

NA

NA

NA

ran\_pars

participant

sd\_\_(Intercept)

0.0000186

NA

NA

NA

ran\_pars

vignette

sd\_\_(Intercept)

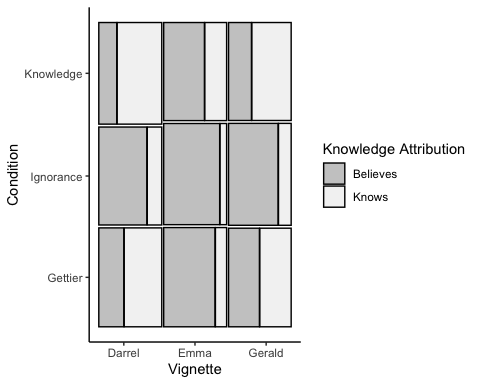
0.0000000

NA

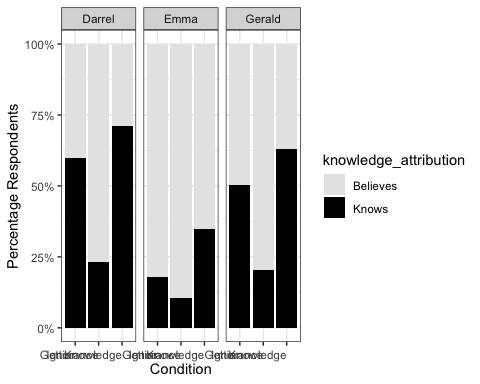
NA

NA

#graph the three way binary   
k.interact <-   
 ggplot(data\_graph) +  
 geom\_mosaic(aes(x = product(know\_vas\_combined, cond, vignette),   
 fill = know\_vas\_combined), color = "black", size = .5) +   
 scale\_fill\_brewer(palette = "Greys", name = "Knowledge Attribution",   
 direction = -1) +   
 scale\_x\_productlist(breaks = c(0.13,.5,.87),  
 labels = c("Darrel", "Emma", "Gerald")) +   
 theme\_classic() +   
 xlab("Vignette") +   
 ylab("Condition") + NULL  
 # ggtitle("Figure X. Rates of Knowledge Attribution by Condition and Vignette")  
  
k.interact



#alternative graph - from Gerit  
  
tknow <- as.data.frame(table(data\_graph$know\_vas\_combined, data\_graph$cond, data\_graph$vignette))  
  
colnames(tknow) <- c("knowledge\_attribution", "Condition", "Vignette", "Frequency")  
  
tknow %>%   
 ggplot(aes(fill = knowledge\_attribution, y = Frequency, x = Condition)) +  
 labs("Figure X. Rates of Knowledge Attribution by Condition and Vignette", fill="knowledge\_attribution") +  
 geom\_bar(position="fill", stat = "identity") +  
 scale\_y\_continuous(labels=scales::percent) +  
 ylab("Percentage Respondents") +  
 scale\_fill\_grey(start = .9, end=0, labels = c("Believes", "Knows")) +  
 theme\_bw()+theme(axis.text.x = element\_text(angle = 0, vjust = 0.5, hjust=1)) +   
 facet\_wrap(~Vignette)



mutate(tidy(k.model.6H), model = "Model 6H") %>%   
 full\_join(  
 mutate(tidy(k.model.6G), model = "Model 6G"),   
 .) %>%   
 full\_join(  
 mutate(tidy(k.model.6F), model = "Model 6F"),  
 .) %>%   
 full\_join(  
 mutate(tidy(k.model.6E), model = "Model 6E"),   
 .) %>%   
 full\_join(  
 mutate(tidy(k.model.6C), model = "Model 6C"),   
 .) %>%   
 full\_join(  
 mutate(tidy(k.model.6B), model = "Model 6B"),   
 .) %>%   
 full\_join(  
 mutate(tidy(k.model.6A), model = "Model 6A"),  
 .) %>%   
 full\_join(  
 mutate(tidy(k.model.6), model = "Model 6"), .) %>%   
 full\_join(  
 mutate(tidy(k.model.5), model = "Model 5"), .) %>%   
 full\_join(  
 mutate(tidy(k.model.4), model = "Model 4"), .) %>%   
 full\_join(  
 mutate(tidy(k.model.3), model = "Model 3"), .) %>%   
 full\_join(  
 mutate(tidy(k.model.2), model = "Model 2"), .) %>%   
 full\_join(  
 mutate(tidy(k.model.1), model = "Model 1"), .) %>%   
 select(., model, effect, group, term, estimate, std.error, statistic, p.value) %>%   
 kable(caption = "Knowledge Attribution Model Results") %>%   
 kable\_styling()

## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("term", "estimate", "std.error", "statistic", "p.value",  
## "model")

Knowledge Attribution Model Results

model

effect

group

term

estimate

std.error

statistic

p.value

Model 1

NA

NA

(Intercept)

-0.4440717

0.0172550

-25.7357786

0.0000000

Model 2

fixed

NA

(Intercept)

-0.4914587

0.3462723

-1.4192838

0.1558163

Model 2

ran\_pars

vignette

sd\_\_(Intercept)

0.5989297

NA

NA

NA

Model 3

fixed

NA

(Intercept)

-0.4914587

0.3462738

-1.4192778

0.1558180

Model 3

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000079

NA

NA

NA

Model 3

ran\_pars

vignette

sd\_\_(Intercept)

0.5989323

NA

NA

NA

Model 4

fixed

NA

(Intercept)

-0.4914587

0.3462705

-1.4192912

0.1558141

Model 4

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0003028

NA

NA

NA

Model 4

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000022

NA

NA

NA

Model 4

ran\_pars

vignette

sd\_\_(Intercept)

0.5989266

NA

NA

NA

Model 5

fixed

NA

(Intercept)

-0.3673361

0.3638228

-1.0096566

0.3126599

Model 5

fixed

NA

compYes

0.0217273

0.0418444

0.5192401

0.6035933

Model 5

fixed

NA

age

0.0034561

0.0018654

1.8527557

0.0639174

Model 5

fixed

NA

gender2male

-0.0685127

0.0399321

-1.7157308

0.0862113

Model 5

fixed

NA

education

-0.0151644

0.0070520

-2.1503593

0.0315268

Model 5

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000030

NA

NA

NA

Model 5

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000918

NA

NA

NA

Model 5

ran\_pars

vignette

sd\_\_(Intercept)

0.6011694

NA

NA

NA

Model 6

fixed

NA

(Intercept)

-0.1838102

0.4040911

-0.4548731

0.6492005

Model 6

fixed

NA

compYes

0.0193640

0.0446327

0.4338527

0.6643954

Model 6

fixed

NA

age

0.0037334

0.0019969

1.8696286

0.0615354

Model 6

fixed

NA

gender2male

-0.0845394

0.0426043

-1.9842938

0.0472231

Model 6

fixed

NA

education

-0.0170686

0.0075284

-2.2672353

0.0233759

Model 6

fixed

NA

condIgnorance

-1.3132223

0.0502375

-26.1402757

0.0000000

Model 6

fixed

NA

condKnowledge

0.6116647

0.0444354

13.7652591

0.0000000

Model 6

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6

ran\_pars

vignette

sd\_\_(Intercept)

0.6688653

NA

NA

NA

Model 6A

fixed

NA

(Intercept)

0.5550196

0.1264092

4.3906588

0.0000113

Model 6A

fixed

NA

compYes

0.0167799

0.0448442

0.3741826

0.7082685

Model 6A

fixed

NA

age

0.0037784

0.0020039

1.8855192

0.0593598

Model 6A

fixed

NA

gender2male

-0.0876738

0.0427943

-2.0487271

0.0404888

Model 6A

fixed

NA

education

-0.0166002

0.0075607

-2.1955787

0.0281221

Model 6A

fixed

NA

condIgnorance

-1.6042917

0.0800233

-20.0478029

0.0000000

Model 6A

fixed

NA

condKnowledge

0.5031718

0.0761979

6.6034827

0.0000000

Model 6A

fixed

NA

vignetteEmma

-1.9304337

0.0844385

-22.8619962

0.0000000

Model 6A

fixed

NA

vignetteGerald

-0.3903202

0.0730638

-5.3421829

0.0000001

Model 6A

fixed

NA

condIgnorance:vignetteEmma

0.9774987

0.1327914

7.3611595

0.0000000

Model 6A

fixed

NA

condKnowledge:vignetteEmma

0.3950256

0.1145277

3.4491693

0.0005623

Model 6A

fixed

NA

condIgnorance:vignetteGerald

0.2231922

0.1137761

1.9616786

0.0497999

Model 6A

fixed

NA

condKnowledge:vignetteGerald

0.0217803

0.1061682

0.2051493

0.8374555

Model 6A

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6A

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000186

NA

NA

NA

Model 6A

ran\_pars

vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6B

fixed

NA

(Intercept)

-0.0743861

0.4058532

-0.1832833

0.8545757

Model 6B

fixed

NA

compYes

-0.0330383

0.0463320

-0.7130761

0.4757986

Model 6B

fixed

NA

age

-0.0001138

0.0022169

-0.0513287

0.9590636

Model 6B

fixed

NA

gender2male

-0.1113136

0.0431663

-2.5787117

0.0099170

Model 6B

fixed

NA

education

-0.0164575

0.0075305

-2.1854445

0.0288563

Model 6B

fixed

NA

condIgnorance

-1.2894721

0.0521660

-24.7186298

0.0000000

Model 6B

fixed

NA

condKnowledge

0.5914056

0.0461318

12.8199170

0.0000000

Model 6B

fixed

NA

turkTRUE

0.3226120

0.1260957

2.5584685

0.0105134

Model 6B

fixed

NA

condIgnorance:turkTRUE

-0.3324922

0.1930846

-1.7220030

0.0850690

Model 6B

fixed

NA

condKnowledge:turkTRUE

0.2986885

0.1735316

1.7212340

0.0852084

Model 6B

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6B

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000024

NA

NA

NA

Model 6B

ran\_pars

vignette

sd\_\_(Intercept)

0.6703837

NA

NA

NA

Model 6C

fixed

NA

(Intercept)

-0.2331619

0.4110771

-0.5671975

0.5705800

Model 6C

fixed

NA

compYes

0.0188339

0.0446380

0.4219245

0.6730801

Model 6C

fixed

NA

age

0.0036811

0.0019978

1.8426411

0.0653814

Model 6C

fixed

NA

gender2male

-0.0844329

0.0426085

-1.9815970

0.0475244

Model 6C

fixed

NA

education

-0.0173380

0.0075339

-2.3013410

0.0213724

Model 6C

fixed

NA

condIgnorance

-1.2653818

0.1351802

-9.3607027

0.0000000

Model 6C

fixed

NA

condKnowledge

0.5910986

0.1189390

4.9697625

0.0000007

Model 6C

fixed

NA

know\_vas\_combined\_sourceVAS

0.0635288

0.0906584

0.7007490

0.4834597

Model 6C

fixed

NA

condIgnorance:know\_vas\_combined\_sourceVAS

-0.0555971

0.1454422

-0.3822626

0.7022666

Model 6C

fixed

NA

condKnowledge:know\_vas\_combined\_sourceVAS

0.0239167

0.1281170

0.1866783

0.8519128

Model 6C

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000018

NA

NA

NA

Model 6C

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000005

NA

NA

NA

Model 6C

ran\_pars

vignette

sd\_\_(Intercept)

0.6689735

NA

NA

NA

Model 6E

fixed

NA

(Intercept)

-0.4781254

0.3577040

-1.3366510

0.1813366

Model 6E

fixed

NA

compYes

0.0244676

0.0490953

0.4983704

0.6182230

Model 6E

fixed

NA

age

0.0027438

0.0021813

1.2578935

0.2084303

Model 6E

fixed

NA

gender2male

-0.1344354

0.0472069

-2.8477931

0.0044024

Model 6E

fixed

NA

education

-0.0169913

0.0083129

-2.0439621

0.0409573

Model 6E

fixed

NA

condIgnorance

-1.0001760

0.0783020

-12.7733145

0.0000000

Model 6E

fixed

NA

condKnowledge

0.3484190

0.0659841

5.2803507

0.0000001

Model 6E

fixed

NA

luck\_vas\_combined

1.0270457

0.0735039

13.9726759

0.0000000

Model 6E

fixed

NA

condIgnorance:luck\_vas\_combined

-1.1288172

0.1127673

-10.0101454

0.0000000

Model 6E

fixed

NA

condKnowledge:luck\_vas\_combined

0.1953581

0.1002016

1.9496502

0.0512178

Model 6E

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6E

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6E

ran\_pars

vignette

sd\_\_(Intercept)

0.5741955

NA

NA

NA

Model 6F

fixed

NA

(Intercept)

0.6703707

0.3812722

1.7582471

0.0787055

Model 6F

fixed

NA

compYes

-0.2247219

0.0444192

-5.0591205

0.0000004

Model 6F

fixed

NA

age

-0.0052593

0.0019654

-2.6759524

0.0074517

Model 6F

fixed

NA

gender2male

0.0941301

0.0420711

2.2374025

0.0252600

Model 6F

fixed

NA

education

0.0056011

0.0074353

0.7533111

0.4512630

Model 6F

fixed

NA

condIgnorance

1.1775177

0.0499634

23.5675933

0.0000000

Model 6F

fixed

NA

condKnowledge

-0.4076711

0.0435151

-9.3684979

0.0000000

Model 6F

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000001

NA

NA

NA

Model 6F

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000011

NA

NA

NA

Model 6F

ran\_pars

vignette

sd\_\_(Intercept)

0.6281604

NA

NA

NA

Model 6G

fixed

NA

(Intercept)

-0.1913248

0.4102302

-0.4663841

0.6409406

Model 6G

fixed

NA

compYes

0.0193490

0.0447266

0.4326054

0.6653014

Model 6G

fixed

NA

age

0.0036964

0.0020017

1.8466406

0.0647993

Model 6G

fixed

NA

gender2male

-0.0870730

0.0427146

-2.0384822

0.0415017

Model 6G

fixed

NA

education

-0.0162490

0.0075470

-2.1530511

0.0313147

Model 6G

fixed

NA

condIgnorance

-1.3186860

0.0503535

-26.1885589

0.0000000

Model 6G

fixed

NA

condKnowledge

0.6140462

0.0445332

13.7884985

0.0000000

Model 6G

fixed

NA

vignette\_orderEGD

-0.0584239

0.0673917

-0.8669308

0.3859799

Model 6G

fixed

NA

vignette\_orderDGE

0.0350471

0.0665816

0.5263780

0.5986256

Model 6G

fixed

NA

vignette\_orderGED

-0.1409815

0.0680785

-2.0708666

0.0383713

Model 6G

fixed

NA

vignette\_orderDEG

-0.0658971

0.0672233

-0.9802723

0.3269517

Model 6G

fixed

NA

vignette\_orderGDE

-0.1742215

0.0673513

-2.5867583

0.0096884

Model 6G

fixed

NA

cond\_orderGKI

-0.1362752

0.0673833

-2.0223895

0.0431361

Model 6G

fixed

NA

cond\_orderIGK

0.1992382

0.0670076

2.9733660

0.0029455

Model 6G

fixed

NA

cond\_orderGIK

0.0196733

0.0671650

0.2929102

0.7695908

Model 6G

fixed

NA

cond\_orderKIG

0.1338955

0.0671760

1.9932049

0.0462390

Model 6G

fixed

NA

cond\_orderIKG

0.1718570

0.0674446

2.5481231

0.0108304

Model 6G

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6G

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000007

NA

NA

NA

Model 6G

ran\_pars

vignette

sd\_\_(Intercept)

0.6715271

NA

NA

NA

Model 6H

fixed

NA

(Intercept)

-0.2439395

0.4181339

-0.5834004

0.5596238

Model 6H

fixed

NA

compYes

0.0192031

0.0447644

0.4289823

0.6679361

Model 6H

fixed

NA

age

0.0037100

0.0020044

1.8508847

0.0641861

Model 6H

fixed

NA

gender2male

-0.0877513

0.0427638

-2.0520006

0.0401696

Model 6H

fixed

NA

education

-0.0162564

0.0075559

-2.1514921

0.0314374

Model 6H

fixed

NA

condIgnorance

-1.1569123

0.1658006

-6.9777332

0.0000000

Model 6H

fixed

NA

condKnowledge

0.6457033

0.1486042

4.3451204

0.0000139

Model 6H

fixed

NA

vignette\_orderEGD

0.0685415

0.1104112

0.6207838

0.5347419

Model 6H

fixed

NA

vignette\_orderDGE

0.0878702

0.1086420

0.8088047

0.4186275

Model 6H

fixed

NA

vignette\_orderGED

-0.0717291

0.1108732

-0.6469469

0.5176663

Model 6H

fixed

NA

vignette\_orderDEG

-0.1231691

0.1099929

-1.1197917

0.2628026

Model 6H

fixed

NA

vignette\_orderGDE

-0.2025115

0.1105884

-1.8312182

0.0670680

Model 6H

fixed

NA

cond\_orderGKI

-0.1361174

0.1097779

-1.2399351

0.2149994

Model 6H

fixed

NA

cond\_orderIGK

0.2873730

0.1087939

2.6414430

0.0082554

Model 6H

fixed

NA

cond\_orderGIK

-0.0006887

0.1095202

-0.0062879

0.9949830

Model 6H

fixed

NA

cond\_orderKIG

0.2719024

0.1096427

2.4798960

0.0131421

Model 6H

fixed

NA

cond\_orderIKG

0.1159230

0.1104895

1.0491767

0.2940968

Model 6H

fixed

NA

condIgnorance:vignette\_orderEGD

-0.2046136

0.1754890

-1.1659620

0.2436298

Model 6H

fixed

NA

condKnowledge:vignette\_orderEGD

-0.2114511

0.1561004

-1.3545839

0.1755502

Model 6H

fixed

NA

condIgnorance:vignette\_orderDGE

-0.0747758

0.1731986

-0.4317344

0.6659345

Model 6H

fixed

NA

condKnowledge:vignette\_orderDGE

-0.0887875

0.1542674

-0.5755424

0.5649245

Model 6H

fixed

NA

condIgnorance:vignette\_orderGED

-0.0591568

0.1783172

-0.3317504

0.7400777

Model 6H

fixed

NA

condKnowledge:vignette\_orderGED

-0.1588843

0.1567770

-1.0134419

0.3108492

Model 6H

fixed

NA

condIgnorance:vignette\_orderDEG

0.1465715

0.1748046

0.8384878

0.4017568

Model 6H

fixed

NA

condKnowledge:vignette\_orderDEG

0.0400427

0.1554705

0.2575580

0.7967481

Model 6H

fixed

NA

condIgnorance:vignette\_orderGDE

-0.0566577

0.1785798

-0.3172684

0.7510399

Model 6H

fixed

NA

condKnowledge:vignette\_orderGDE

0.1158792

0.1558071

0.7437355

0.4570365

Model 6H

fixed

NA

condIgnorance:cond\_orderGKI

-0.0840584

0.1767048

-0.4756995

0.6342885

Model 6H

fixed

NA

condKnowledge:cond\_orderGKI

0.0614584

0.1540798

0.3988739

0.6899861

Model 6H

fixed

NA

condIgnorance:cond\_orderIGK

-0.1761919

0.1735113

-1.0154496

0.3098916

Model 6H

fixed

NA

condKnowledge:cond\_orderIGK

-0.1254365

0.1538993

-0.8150552

0.4150407

Model 6H

fixed

NA

condIgnorance:cond\_orderGIK

-0.1715086

0.1783025

-0.9618968

0.3361014

Model 6H

fixed

NA

condKnowledge:cond\_orderGIK

0.1826659

0.1542085

1.1845383

0.2362000

Model 6H

fixed

NA

condIgnorance:cond\_orderKIG

-0.2803621

0.1760733

-1.5923033

0.1113166

Model 6H

fixed

NA

condKnowledge:cond\_orderKIG

-0.1686158

0.1546747

-1.0901314

0.2756553

Model 6H

fixed

NA

condIgnorance:cond\_orderIKG

-0.0169973

0.1754102

-0.0969004

0.9228055

Model 6H

fixed

NA

condKnowledge:cond\_orderIKG

0.1609219

0.1561290

1.0306984

0.3026823

Model 6H

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6H

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000002

NA

NA

NA

Model 6H

ran\_pars

vignette

sd\_\_(Intercept)

0.6722918

NA

NA

NA

*Reasonableness attributions*

As a secondary dependent measurement, judgments of reasonableness were predicted in a series of logistic regression models paralleling those for knowledge attributions. Compared to a baseline intercept-only model (Model 1, AIC = 7343.35), a model with a random intercept for vignette (Model 2, AIC = 7286.55) explained more variance. The likelihood of the protagonist being judged as reasonable varied by vignette (Pseudo-*R*2 = 0 to 0.02); although, overall participants were far more likely to respond that the protagonist was reasonable than unreasonable in all three vignettes. Collapsing across conditions, participants were more likely to judge Emma as unreasonable than Gerald. Participants were more likely to judge Gerald as unreasonable than Darrel (See Table X).

table(r.model.2@frame$reason\_vas\_combined,   
 r.model.2@frame$vignette) %>%   
 kable(caption = "Reasonableness by Vignette") %>%   
 kable\_styling()

Reasonableness by Vignette

Darrel

Emma

Gerald

0

237

447

337

1

4493

4269

4411

A model with a random intercept for vignette nested within participant (Model 3, AIC = 7288.56) explained similar amounts of variance as Model 2 (Pseudo-*R*2 = 0 to 0.02). The model with a random intercept for vignette nested in participant nested in data collection site (Model 4, AIC = 7290.55) did not explain more variance than previous models (Pseudo-*R*2 = 0 to 0.02). In Model 5, covariates were added as fixed effects (AIC = 7144.1). Relative to Model 4, this model was more useful in explaining variance in judgements of reasonableness (Pseudo-*R*2 = 0.008 - 0.037). Compensation was associated with reasonableness. That is, participants who were compensated were more likely to judge the protagonist as reasonable. Female participants were more likely to select reasonableness over male participants. Education was also associated with reasonableness. As the participant’s years of education decreased, the likelihood that they would judge the protagonist as reasonable increased.

Finally, we estimated a model including knowledge condition as a fixed effect (Model 6, AIC = 7047.13). This model performed better than Model 5 and revealed an effect of condition on reasonableness judgment (Pseudo-*R*2 = 0.01 to 0.05). Participants were more likely to judge the protagonist in the knowledge condition vignette as reasonable than the protagonists in the other two conditions (see Table X). Protagonists in the ignorance condition vignette were less likely to be judged as reasonable than protagonists in the knowledge and Gettier condition vignettes.

table(r.model.6@frame$reason\_vas\_combined,   
 r.model.6@frame$cond) %>%   
 kable(caption = "Reasonableness by Condition") %>%   
 kable\_styling()

Reasonableness by Condition

Gettier

Ignorance

Knowledge

0

323

467

217

1

4328

4192

4447

*Does the effect of condition differ by vignette?*

tidy(r.model.6A) %>%   
 clean\_variable\_names() %>%   
 kable(caption = "Interaction of Condition and Vignette on Reasonableness") %>%   
 kable\_styling()

Interaction of Condition and Vignette on Reasonableness

effect

group

term

estimate

std.error

statistic

p.value

fixed

NA

(Intercept)

2.4251103

0.2229400

10.8778606

0.0000000

fixed

NA

compensation

0.2489033

0.0725555

3.4305237

0.0006024

fixed

NA

age

-0.0050065

0.0032180

-1.5557632

0.1197644

fixed

NA

gender

-0.1818856

0.0706659

-2.5738806

0.0100565

fixed

NA

education

0.0585941

0.0119364

4.9088712

0.0000009

fixed

NA

condIgnorance

-0.8637051

0.1605822

-5.3785861

0.0000001

fixed

NA

condKnowledge

0.3908751

0.2046418

1.9100452

0.0561274

fixed

NA

vignetteEmma

-1.0995648

0.1556147

-7.0659426

0.0000000

fixed

NA

vignetteGerald

-0.5223064

0.1682765

-3.1038588

0.0019101

fixed

NA

condIgnorance:vignetteEmma

0.7411950

0.1968864

3.7645824

0.0001668

fixed

NA

condKnowledge:vignetteEmma

0.2566166

0.2458387

1.0438415

0.2965587

fixed

NA

condIgnorance:vignetteGerald

0.4243788

0.2100619

2.0202564

0.0433568

fixed

NA

condKnowledge:vignetteGerald

-0.2371810

0.2558282

-0.9271102

0.3538693

ran\_pars

site

sd\_\_(Intercept)

0.0001301

NA

NA

NA

ran\_pars

participant

sd\_\_(Intercept)

0.0460785

NA

NA

NA

ran\_pars

vignette

sd\_\_(Intercept)

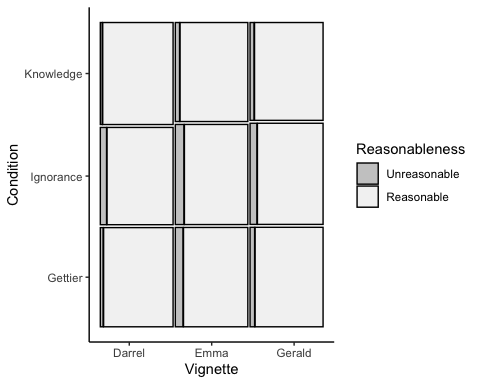
0.0000000

NA

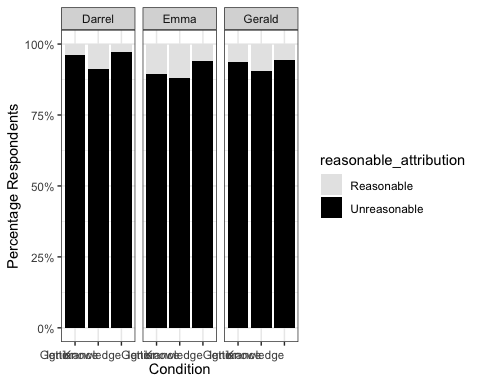
NA

NA

#graph the three way binary   
r.interact <- ggplot(r.data\_graph) +  
 geom\_mosaic(aes(x = product(reason\_vas\_combined, cond, vignette),   
 fill = reason\_vas\_combined), color = "black", size = .5) +   
 scale\_fill\_brewer(palette = "Greys", name = "Reasonableness",   
 direction = -1) +   
 scale\_x\_productlist(breaks = c(0.13,.5,.87),  
 labels = c("Darrel", "Emma", "Gerald")) +   
 theme\_classic() +   
 xlab("Vignette") +   
 ylab("Condition") + NULL  
 # ggtitle("Figure X. Reasonableness Judgments by Condition and Vignette")  
  
r.interact



#alternative graph - from Gerit  
  
rgraph <- as.data.frame(table(r.data\_graph$reason\_vas\_combined, r.data\_graph$cond, r.data\_graph$vignette))  
  
colnames(rgraph) <- c("reasonableness\_attribution", "Condition", "Vignette", "Frequency")  
  
rgraph %>%   
 ggplot(aes(fill = reasonableness\_attribution, y = Frequency, x = Condition)) +  
 labs("Figure X. Rates of Reasonableness Attribution by Condition and Vignette", fill="reasonable\_attribution") +  
 geom\_bar(position="fill", stat = "identity") +  
 scale\_y\_continuous(labels=scales::percent) +  
 ylab("Percentage Respondents") +  
 scale\_fill\_grey(start = .9, end=0, labels = c("Reasonable", "Unreasonable")) +  
 theme\_bw()+theme(axis.text.x = element\_text(angle = 0, vjust = 0.5, hjust=1)) +   
 facet\_wrap(~Vignette)



r.dcond\_effect <- stats::chisq.test(data\_graph$know\_vas\_combined[data\_graph$vignette == "Darrel"],   
 data\_graph$cond[data\_graph$vignette == "Darrel"]) %>%   
 tidy()  
  
r.dcond\_GI <- prop.test(t(r.d\_table[2:1, 1:2])) %>%   
 tidy()  
  
r.dcond\_GI\_v <- v.chi.sq(x2 = prop.test(t(r.d\_table[2:1, 1:2]))$statistic,   
 n = sum(t(r.d\_table[2:1, 1:2])),  
 r = 2, c = 2)   
  
r.dcond\_GK <- prop.test(t(r.d\_table[2:1, c(1,3)])) %>%   
 tidy()  
  
r.dcond\_GK\_v <- v.chi.sq(x2 = prop.test(t(r.d\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(r.d\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

## Warning: The size of the effect combined with the degrees of freedom is too  
## small to determine a lower confidence limit for the 'alpha.lower' (or the (1/2)  
## (1-'conf.level') symmetric) value specified (set to zero).

r.econd\_effect <- stats::chisq.test(r.data\_graph$reason\_vas\_combined[r.data\_graph$vignette == "Emma"],   
 r.data\_graph$cond[r.data\_graph$vignette == "Emma"]) %>%   
 tidy()  
  
r.econd\_GI <- prop.test(t(r.e\_table[2:1, 1:2])) %>%   
 tidy()  
  
r.econd\_GI\_v <- v.chi.sq(x2 = prop.test(t(r.e\_table[2:1, 1:2]))$statistic,   
 n = sum(t(r.e\_table[2:1, 1:2])),  
 r = 2, c = 2)

## Warning: The size of the effect combined with the degrees of freedom is too  
## small to determine a lower confidence limit for the 'alpha.lower' (or the (1/2)  
## (1-'conf.level') symmetric) value specified (set to zero).

r.econd\_GK <- prop.test(t(r.e\_table[2:1, c(1,3)])) %>%   
 tidy()  
  
r.econd\_GK\_v <- v.chi.sq(x2 = prop.test(t(r.e\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(r.e\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

r.gcond\_effect <- stats::chisq.test(r.data\_graph$reason\_vas\_combined[r.data\_graph$vignette == "Gerald"],   
 r.data\_graph$cond[r.data\_graph$vignette == "Gerald"]) %>%   
 tidy()  
  
r.gcond\_GI <- prop.test(t(r.g\_table[2:1, 1:2])) %>%   
 tidy()  
  
r.gcond\_GI\_v <- v.chi.sq(x2 = prop.test(t(r.g\_table[2:1, 1:2]))$statistic,   
 n = sum(t(r.g\_table[2:1, 1:2])),  
 r = 2, c = 2)   
  
r.gcond\_GK <- prop.test(t(r.g\_table[2:1, c(1,3)])) %>%   
 tidy()  
  
r.gcond\_GK\_v <- v.chi.sq(x2 = prop.test(t(r.g\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(r.g\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

## Warning: The size of the effect combined with the degrees of freedom is too  
## small to determine a lower confidence limit for the 'alpha.lower' (or the (1/2)  
## (1-'conf.level') symmetric) value specified (set to zero).

To test whether the effect of condition on reasonableness judgments varied by vignette, a model was estimated that included an interaction between vignette and condition (Model 6A, AIC = 7025.8). This model explained more variance than the model without the interaction term. As shown in Figure X, although the general pattern is the same for all vignettes, the magnitudes of differences vary by vignette (Pseudo-*R*2 = 0.02% to 0.08).

The likelihood that participants judged the protagonist as reasonable varied by condition in response to the Darrel vignette, (2) = 781, p < .001, Emma vignette (2) = 36.36, p = 0, and Gerald vignette (2) = 21.1, p = 0. Participants were more likely to judge Darrel to be reasonable in the Gettier condition vignette (0.96) than in the ignorance condition (0.91; Cramer’s = .10, 95% CI [.06, .13], (1) = 28.84, < .001), but reasonable judgments did not differ between participants responding to the knowledge and Gettier conditions (Cramer’s = .03, 95% CI [.02, .07], (1) = 3.44, = .064). The same pattern of results appeared in response to the Gerald vignette; participants were more likely to judge Gerald as reasonable when responding to the Gettier condition vignette (0.94) as opposed to the ignorance condition (0.91; Cramer’s = .06, 95% CI [.03, .09], (1) = 10.49, = .001), but the knowledge and Gettier vignettes produced similar rates of reasonableness judgments, (0.94; Cramer’s = .02, 95% CI [.02, .05], (1) = 0.77, = .381).

The condition by vignette interaction in predicting judgements of reasonableness appears to have emerged because of the condition differences produced by the Emma vignette. While participants were equally likely to judge Emma as reasonable in the Gettier and ignorance conditions, (Cramer’s = .02, 95% CI [.02, .06], (1) = 1.12, = .291), participants were more likely to judge Emma as reasonable in response to the knowledge condition vignette (0.94) than in response to the the Gettier condition vignette (0.89; Cramer’s = .09, 95% CI [.05, .12], (1) = 22.44, < .001). Thus, condition differences were found between the Gettier and ignorance versions of the Darrel and Gerald vignettes, but not the Emma vignette, and between the Gettier and knowledge versions of the Emma vignette, but not the Darrel and Gerald vignettes.

mutate(tidy(r.model.6C), model = "Model 6C") %>%   
 full\_join(  
 mutate(tidy(r.model.6B), model = "Model 6B"),   
 .) %>%   
 full\_join(  
 mutate(tidy(r.model.6A), model = "Model 6A"),  
 .) %>%   
 full\_join(  
 mutate(tidy(r.model.6), model = "Model 6"), .) %>%   
 full\_join(  
 mutate(tidy(r.model.5), model = "Model 5"), .) %>%   
 full\_join(  
 mutate(tidy(r.model.4), model = "Model 4"), .) %>%   
 full\_join(  
 mutate(tidy(r.model.3), model = "Model 3"), .) %>%   
 full\_join(  
 mutate(tidy(r.model.2), model = "Model 2"), .) %>%   
 full\_join(  
 mutate(tidy(r.model.1), model = "Model 1"), .) %>%   
 select(., model, effect, group, term, estimate, std.error, statistic, p.value) %>%   
 kable(caption = "Reasonableness Attribution Model Results") %>%   
 kable\_styling()

## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("term", "estimate", "std.error", "statistic", "p.value",  
## "model")

Reasonableness Attribution Model Results

model

effect

group

term

estimate

std.error

statistic

p.value

Model 1

NA

NA

(Intercept)

2.5573867

0.0324860

78.7226670

0.0000000

Model 2

fixed

NA

(Intercept)

2.5873781

0.1616973

16.0013715

0.0000000

Model 2

ran\_pars

vignette

sd\_\_(Intercept)

0.2740345

NA

NA

NA

Model 3

fixed

NA

(Intercept)

2.5873777

0.1616883

16.0022539

0.0000000

Model 3

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000078

NA

NA

NA

Model 3

ran\_pars

vignette

sd\_\_(Intercept)

0.2740205

NA

NA

NA

Model 4

fixed

NA

(Intercept)

2.5873778

0.1616901

16.0020837

0.0000000

Model 4

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 4

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0001354

NA

NA

NA

Model 4

ran\_pars

vignette

sd\_\_(Intercept)

0.2740214

NA

NA

NA

Model 5

fixed

NA

(Intercept)

1.8072344

0.2439704

7.4075962

0.0000000

Model 5

fixed

NA

compYes

0.2479975

0.0721679

3.4363959

0.0005895

Model 5

fixed

NA

age

-0.0050339

0.0031939

-1.5760992

0.1150029

Model 5

fixed

NA

gender2male

-0.1794396

0.0703104

-2.5521052

0.0107074

Model 5

fixed

NA

education

0.0574914

0.0118537

4.8500873

0.0000012

Model 5

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000698

NA

NA

NA

Model 5

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 5

ran\_pars

vignette

sd\_\_(Intercept)

0.2783242

NA

NA

NA

Model 6

fixed

NA

(Intercept)

1.8398187

0.2489283

7.3909591

0.0000000

Model 6

fixed

NA

compYes

0.2515650

0.0725041

3.4696649

0.0005211

Model 6

fixed

NA

age

-0.0049681

0.0032090

-1.5481727

0.1215807

Model 6

fixed

NA

gender2male

-0.1811528

0.0706323

-2.5647309

0.0103256

Model 6

fixed

NA

education

0.0577126

0.0119259

4.8392623

0.0000013

Model 6

fixed

NA

condIgnorance

-0.4002706

0.0759400

-5.2708777

0.0000001

Model 6

fixed

NA

condKnowledge

0.4256132

0.0906709

4.6940423

0.0000027

Model 6

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0226841

NA

NA

NA

Model 6

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0908005

NA

NA

NA

Model 6

ran\_pars

vignette

sd\_\_(Intercept)

0.2768698

NA

NA

NA

Model 6A

fixed

NA

(Intercept)

2.4251103

0.2229400

10.8778606

0.0000000

Model 6A

fixed

NA

compYes

0.2489033

0.0725555

3.4305237

0.0006024

Model 6A

fixed

NA

age

-0.0050065

0.0032180

-1.5557632

0.1197644

Model 6A

fixed

NA

gender2male

-0.1818856

0.0706659

-2.5738806

0.0100565

Model 6A

fixed

NA

education

0.0585941

0.0119364

4.9088712

0.0000009

Model 6A

fixed

NA

condIgnorance

-0.8637051

0.1605822

-5.3785861

0.0000001

Model 6A

fixed

NA

condKnowledge

0.3908751

0.2046418

1.9100452

0.0561274

Model 6A

fixed

NA

vignetteEmma

-1.0995648

0.1556147

-7.0659426

0.0000000

Model 6A

fixed

NA

vignetteGerald

-0.5223064

0.1682765

-3.1038588

0.0019101

Model 6A

fixed

NA

condIgnorance:vignetteEmma

0.7411950

0.1968864

3.7645824

0.0001668

Model 6A

fixed

NA

condKnowledge:vignetteEmma

0.2566166

0.2458387

1.0438415

0.2965587

Model 6A

fixed

NA

condIgnorance:vignetteGerald

0.4243788

0.2100619

2.0202564

0.0433568

Model 6A

fixed

NA

condKnowledge:vignetteGerald

-0.2371810

0.2558282

-0.9271102

0.3538693

Model 6A

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0001301

NA

NA

NA

Model 6A

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0460785

NA

NA

NA

Model 6A

ran\_pars

vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6B

fixed

NA

(Intercept)

2.0738657

0.2515542

8.2442097

0.0000000

Model 6B

fixed

NA

compYes

0.1317760

0.0757080

1.7405816

0.0817570

Model 6B

fixed

NA

age

-0.0124663

0.0033764

-3.6922327

0.0002223

Model 6B

fixed

NA

gender2male

-0.2323798

0.0710885

-3.2688804

0.0010797

Model 6B

fixed

NA

education

0.0568432

0.0117246

4.8482044

0.0000012

Model 6B

fixed

NA

condIgnorance

-0.3783433

0.0774512

-4.8849274

0.0000010

Model 6B

fixed

NA

condKnowledge

0.4401026

0.0923144

4.7674328

0.0000019

Model 6B

fixed

NA

turkTRUE

1.3850376

0.3688729

3.7547833

0.0001735

Model 6B

fixed

NA

condIgnorance:turkTRUE

-0.6553836

0.4279269

-1.5315316

0.1256381

Model 6B

fixed

NA

condKnowledge:turkTRUE

-0.4437035

0.5148000

-0.8618950

0.3887453

Model 6B

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0486724

NA

NA

NA

Model 6B

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0003045

NA

NA

NA

Model 6B

ran\_pars

vignette

sd\_\_(Intercept)

0.2772821

NA

NA

NA

Model 6C

fixed

NA

(Intercept)

1.5691162

0.2767745

5.6692946

0.0000000

Model 6C

fixed

NA

compYes

0.2461394

0.0726686

3.3871468

0.0007062

Model 6C

fixed

NA

age

-0.0053531

0.0032240

-1.6604224

0.0968295

Model 6C

fixed

NA

gender2male

-0.1802269

0.0706571

-2.5507265

0.0107499

Model 6C

fixed

NA

education

0.0566898

0.0120213

4.7157818

0.0000024

Model 6C

fixed

NA

condIgnorance

-0.3422345

0.1827418

-1.8727763

0.0610993

Model 6C

fixed

NA

condKnowledge

0.4773118

0.2170240

2.1993500

0.0278530

Model 6C

fixed

NA

reason\_vas\_combined\_sourceVAS

0.3536665

0.1516900

2.3315084

0.0197266

Model 6C

fixed

NA

condIgnorance:reason\_vas\_combined\_sourceVAS

-0.0708420

0.2009448

-0.3525445

0.7244300

Model 6C

fixed

NA

condKnowledge:reason\_vas\_combined\_sourceVAS

-0.0624726

0.2388977

-0.2615035

0.7937043

Model 6C

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0856972

NA

NA

NA

Model 6C

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0697774

NA

NA

NA

Model 6C

ran\_pars

vignette

sd\_\_(Intercept)

0.2771243

NA

NA

NA

**Participant Recruitment**

Because data was collected from MTurk workers as well as participants recruited from individual labs, we explored whether participant recruitment moderated the effect of condition on knowledge attributions and the other two dependent measures. Though Model 6B (AIC = 15850.16) was superior to Model 6, the interaction term was not a significant predictor of knowledge attributions (Pseudo-*R*2 =0 - 0.01). Next, we estimated the same model (Model 6B) in predicting judgments of reasonableness (AIC = 7017.37). While this model performed better than Model 6, the interaction between condition and recruitment type was not significant (Pseudo-*R*2 =0 - 0.01).

**Exploratory Analyses**

In addition to the hypotheses and research questions outlined in the approved protocol, we conducted additional exploratory analyses to examine two additional research questions and assess the influence of original measurement characteristics (binary vs. continuous).

*“Direct” Replication Analysis*

k.d\_table\_rep <- table(darrel\_DF$know\_vas\_combined, darrel\_DF$cond)  
  
k.dcond\_GI\_rep <- prop.test(t(k.d\_table\_rep[2:1, 1:2])) %>%   
 tidy()  
  
k.dcond\_GI\_v\_rep <- v.chi.sq(x2 = prop.test(t(k.d\_table\_rep[2:1, 1:2]))$statistic,   
 n = sum(t(k.d\_table\_rep[2:1, 1:2])),  
 r = 2, c = 2)

## Warning: The size of the effect combined with the degrees of freedom is too  
## small to determine a lower confidence limit for the 'alpha.lower' (or the (1/2)  
## (1-'conf.level') symmetric) value specified (set to zero).

k.dcond\_GK\_rep <- prop.test(t(k.d\_table\_rep[2:1, c(1,3)])) %>%   
 tidy()  
  
k.dcond\_GK\_v\_rep <- v.chi.sq(x2 = prop.test(t(k.d\_table\_rep[2:1, c(1,3)]))$statistic,   
 n = sum(t(k.d\_table\_rep[2:1, c(1,3)])),  
 r = 2, c = 2)

As previously explained, the design of our study substantially differed from that of Turri et al. (2015, Experiment 1). Rather than encountering one of three conditions of the Darrel/squirrel vignette, our participants viewed three conditions matched with three vignettes in a within-subjects design. Perhaps our observation of a Gettier intuition effect, which was not found in the original experiment, can be explained by these methodological changes. To explore this possibility, we compared the knowledge attribution rates of participants who viewed the Darrel vignette first in an analysis devised to closely approximate Turri et al. ’s original test. 2002 participants correctly answered the Darrel vignette when it appeared as the first vignette. The difference in conditions was significant, (2) = 31.14, < .001, = .14, 95% CI [.09, .18]. No significant difference in Gettier ( = 0.59) and Ignorance ( = 0.64) were found, (1) = 2.14, = .144, = .05, 95% CI [.03, .12]. Knowledge conditions ( = 0.72) selected knows more Gettier conditions, (1) = 30.48, < .001, = .14, 95% CI [.10, .20].

*Luck Attributions*

Attributions of luck were predicted in a series of multilevel logistic regressions models. These models were fit in the same fashion as the models focused on the two dependent variables, with one notable difference: observations where the participant did not correctly answer the first part of our two-part luck attribution measure were excluded. That is, the luck versus ability attributions that followed incorrect identification responses were excluded from analyses (*n* = 952; 6.58%).

Compared to the baseline intercept-only model (Model 1, AIC = 11269.61), a model with a random intercept for vignette (Model 2, AIC = 10613.78) explained more variance. The likelihood that outcomes were attributed to luck varied according to vignette (Pseudo-*R*2 = 0.08 to 0.09). While the Darrel vignette produced more attributions to ability than luck, the Emma vignette produced more attributions to luck than ability.

table(l.model.2@frame$luck\_vas\_combined,   
 l.model.2@frame$vignette) %>%   
 kable(caption = "Luck by Vignette") %>%   
 kable\_styling()

Luck by Vignette

Darrel

Emma

Gerald

0

1434

2910

2207

1

2864

1401

2010

A model with a random intercept for vignette nested within participant (Model 3, AIC = 16773.3) explained similar amounts of variance as the previous model (Pseudo-*R*2 = 0.08 - 0.09). The model with a random intercept for vignette nested in participant nested in data collection site (Model 4, AIC = 16775.3) accounted for the same proportion of random variance (Pseudo-*R*2 = 0.08 - 0.09). Next, covariates were added to the model as fixed effects (Model 5, AIC = 16489.6). Relative to Model 4, Model 5 explained more variance in luck attributions (Pseudo-*R*2 = 0.08 - 0.1). Years of education emerged as a significant predictor; as participant education increased, the likelihood that they would attribute the protagonist’s success to ability decreased (-2.9% per year of education). Participants who were compensated were also more likely to pick luck, as well as younger individuals (e.g., increases in age were associated with a higher likelihood of picking ability).

tidy(l.model.5) %>%   
 clean\_variable\_names() %>%   
 kable(caption = "Explaining Variance in Luck Attributions with Random Intercepts for Vignette/Participant/Site and Covariates") %>%   
 kable\_styling()

Explaining Variance in Luck Attributions with Random Intercepts for Vignette/Participant/Site and Covariates

effect

group

term

estimate

std.error

statistic

p.value

fixed

NA

(Intercept)

0.4036010

0.3558809

1.1340900

0.2567568

fixed

NA

compensation

-0.1777094

0.0428462

-4.1476116

0.0000336

fixed

NA

age

0.0035660

0.0019174

1.8598358

0.0629088

fixed

NA

gender

0.0074075

0.0411403

0.1800541

0.8571101

fixed

NA

education

-0.0294521

0.0073157

-4.0258534

0.0000568

ran\_pars

site

sd\_\_(Intercept)

0.0000000

NA

NA

NA

ran\_pars

participant

sd\_\_(Intercept)

0.0000000

NA

NA

NA

ran\_pars

vignette

sd\_\_(Intercept)

0.5846120

NA

NA

NA

Finally, we estimated a model including condition as a fixed effect (Model 6, AIC = 15896.17). This model performed better than the previous models; the likelihood of luck attributions differed according to condition (Pseudo-*R*2 = 0.05 to 0.06). Participants were more likely to attribute the outcome to luck in the Gettier condition than in the other two conditions (see Table X). In response to both the knowledge condition and the ignorance condition, participants were more likely to attribute outcomes to the protagonist’s ability than to luck, but they were more likely to make luck attributions than ability attributions in response to the Gettier condition vignette.

tidy(l.model.6) %>%   
 clean\_variable\_names() %>%   
 kable(caption = "Effect of Condition on Luck Attributions") %>%   
 kable\_styling()

Effect of Condition on Luck Attributions

effect

group

term

estimate

std.error

statistic

p.value

fixed

NA

(Intercept)

-0.2463155

0.3722464

-0.6617002

0.5081634

fixed

NA

compensation

-0.1806945

0.0439764

-4.1088941

0.0000398

fixed

NA

age

0.0041529

0.0019673

2.1109937

0.0347729

fixed

NA

gender

0.0072667

0.0421847

0.1722583

0.8632345

fixed

NA

education

-0.0310962

0.0074846

-4.1546683

0.0000326

fixed

NA

condIgnorance

1.0277337

0.0474143

21.6756170

0.0000000

fixed

NA

condKnowledge

0.9356795

0.0471059

19.8633078

0.0000000

ran\_pars

site

sd\_\_(Intercept)

0.0991901

NA

NA

NA

ran\_pars

participant

sd\_\_(Intercept)

0.0662757

NA

NA

NA

ran\_pars

vignette

sd\_\_(Intercept)

0.6111192

NA

NA

NA

table(l.model.6@frame$cond,   
 str\_replace(  
 str\_replace(l.model.6@frame$luck\_vas\_combined, "0", "Luck"),  
 "1", "Ability"))

##   
## Ability Luck  
## Gettier 1444 2779  
## Ignorance 2393 1784  
## Knowledge 2352 1888

*Does the effect of condition on luck attributions differ by vignette?*

To better understand whether the effect of condition on luck attributions varied as a function of vignette, we estimated a model including an interaction between vignette and condition (Model 6A, AIC = 15458.37). This model explained more variance (Pseudo-*R*2 = 0.2 - 0.23) than Model 6. As shown in Figure X, each vignette demonstrated a different pattern of effects.

tidy(l.model.6A) %>%   
 clean\_variable\_names() %>%   
 kable(caption = "Interaction of Condition and Vignette on Luck Attributions") %>%   
 kable\_styling()

Interaction of Condition and Vignette on Luck Attributions

effect

group

term

estimate

std.error

statistic

p.value

fixed

NA

(Intercept)

0.9073043

0.1285775

7.0564806

0.0000000

fixed

NA

compensation

-0.1953080

0.0446956

-4.3697376

0.0000124

fixed

NA

age

0.0041522

0.0020023

2.0737491

0.0381026

fixed

NA

gender

-0.0042108

0.0427246

-0.0985578

0.9214894

fixed

NA

education

-0.0310816

0.0076074

-4.0856980

0.0000439

fixed

NA

condIgnorance

0.0839290

0.0784178

1.0702794

0.2844936

fixed

NA

condKnowledge

0.7235533

0.0824618

8.7744067

0.0000000

fixed

NA

vignetteEmma

-2.8510151

0.1106195

-25.7731706

0.0000000

fixed

NA

vignetteGerald

-1.0704438

0.0788193

-13.5809879

0.0000000

fixed

NA

condIgnorance:vignetteEmma

2.5277977

0.1350603

18.7160679

0.0000000

fixed

NA

condKnowledge:vignetteEmma

1.0673419

0.1382340

7.7212696

0.0000000

fixed

NA

condIgnorance:vignetteGerald

0.7232640

0.1105027

6.5452171

0.0000000

fixed

NA

condKnowledge:vignetteGerald

0.0949037

0.1137941

0.8339948

0.4042839

ran\_pars

site

sd\_\_(Intercept)

0.0215039

NA

NA

NA

ran\_pars

participant

sd\_\_(Intercept)

0.0395744

NA

NA

NA

ran\_pars

vignette

sd\_\_(Intercept)

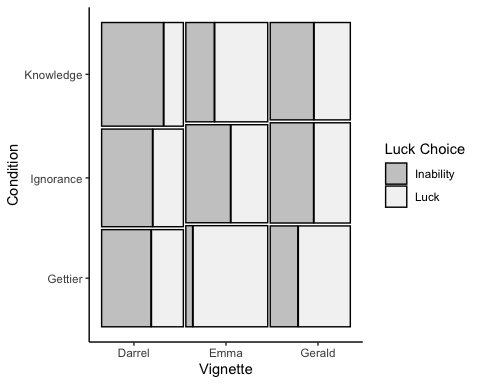
0.0000000

NA

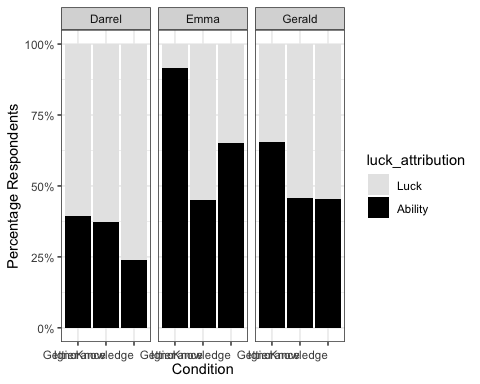
NA

NA

#graph the three way binary   
l.data\_graph$luck\_vas\_combined <- factor(l.data\_graph$luck\_vas\_combined,   
 levels = c("Ability", "Luck"),  
 labels = c("Inability", "Luck"))  
l.interact <- ggplot(l.data\_graph) +  
 geom\_mosaic(aes(x = product(luck\_vas\_combined, cond, vignette),   
 fill = luck\_vas\_combined), color = "black", size = .5) +   
 scale\_fill\_brewer(palette = "Greys", name = "Luck Choice",   
 direction = -1) +   
 scale\_x\_productlist(breaks = c(0.13,.5,.87),  
 labels = c("Darrel", "Emma", "Gerald")) +   
 theme\_classic() +   
 xlab("Vignette") +   
 ylab("Condition") +  
 NULL  
 #ggtitle("Figure X. Rates of Luck Attribution by Condition and Vignette")  
  
l.interact



#alternative graph - from Gerit  
lgraph <- as.data.frame(table(l.data\_graph$luck\_vas\_combined, l.data\_graph$cond, l.data\_graph$vignette))  
  
colnames(lgraph) <- c("luck\_attribution", "Condition", "Vignette", "Frequency")  
  
lgraph %>%   
 ggplot(aes(fill = luck\_attribution, y = Frequency, x = Condition)) +  
 labs("Figure X. Rates of Luck Attribution by Condition and Vignette", fill="luck\_attribution") +  
 geom\_bar(position="fill", stat = "identity") +  
 scale\_y\_continuous(labels=scales::percent) +  
 ylab("Percentage Respondents") +  
 scale\_fill\_grey(start = .9, end=0, labels = c("Luck", "Ability")) +  
 theme\_bw()+theme(axis.text.x = element\_text(angle = 0, vjust = 0.5, hjust=1)) +   
 facet\_wrap(~Vignette)



l.dcond\_effect <- stats::chisq.test(l.data\_graph$luck\_vas\_combined[l.data\_graph$vignette == "Darrel"],   
 l.data\_graph$cond[l.data\_graph$vignette == "Darrel"]) %>%   
 tidy()  
  
l.dcond\_GI <- prop.test(t(l.d\_table[1:2, 1:2])) %>%   
 tidy()  
  
l.dcond\_GI\_v <- v.chi.sq(x2 = prop.test(t(l.d\_table[1:2, 1:2]))$statistic,   
 n = sum(t(l.d\_table[1:2, 1:2])),  
 r = 2, c = 2)

## Warning: The size of the effect combined with the degrees of freedom is too  
## small to determine a lower confidence limit for the 'alpha.lower' (or the (1/2)  
## (1-'conf.level') symmetric) value specified (set to zero).

l.dcond\_GK <- prop.test(t(l.d\_table[1:2, c(1,3)])) %>%   
 tidy()  
  
l.dcond\_GK\_v <- v.chi.sq(x2 = prop.test(t(l.d\_table[1:2, c(1,3)]))$statistic,   
 n = sum(t(l.d\_table[1:2, c(1,3)])),  
 r = 2, c = 2)

l.econd\_effect <- stats::chisq.test(l.data\_graph$luck\_vas\_combined[l.data\_graph$vignette == "Emma"],   
 l.data\_graph$cond[l.data\_graph$vignette == "Emma"]) %>%   
 tidy()  
  
l.econd\_GI <- prop.test(t(l.e\_table[1:2, 1:2])) %>%   
 tidy()  
  
l.econd\_GI\_v <- v.chi.sq(x2 = prop.test(t(l.e\_table[1:2, 1:2]))$statistic,   
 n = sum(t(l.e\_table[1:2, 1:2])),  
 r = 2, c = 2)   
  
l.econd\_GK <- prop.test(t(l.e\_table[1:2, c(1,3)])) %>%   
 tidy()  
  
l.econd\_GK\_v <- v.chi.sq(x2 = prop.test(t(l.e\_table[1:2, c(1,3)]))$statistic,   
 n = sum(t(l.e\_table[1:2, c(1,3)])),  
 r = 2, c = 2)

l.gcond\_effect <- stats::chisq.test(data\_graph$know\_vas\_combined[data\_graph$vignette == "Gerald"],   
 data\_graph$cond[data\_graph$vignette == "Gerald"]) %>%   
 tidy()  
  
l.gcond\_GI <- prop.test(t(l.e\_table[1:2, 1:2])) %>%   
 tidy()  
  
l.gcond\_GI\_v <- v.chi.sq(x2 = prop.test(t(l.g\_table[1:2, 1:2]))$statistic,   
 n = sum(t(l.e\_table[1:2, 1:2])),  
 r = 2, c = 2)   
  
l.gcond\_GK <- prop.test(t(l.e\_table[1:2, c(1,3)])) %>%   
 tidy()  
  
l.gcond\_GK\_v <- v.chi.sq(x2 = prop.test(t(l.g\_table[1:2, c(1,3)]))$statistic,   
 n = sum(t(l.e\_table[1:2, c(1,3)])),  
 r = 2, c = 2)

The likelihood that participants attributed the outcome in the Darrel vignette to luck differed according to condition, (2) = 90.79, p < .001. While participants were equally likely to report luck attributions in response to the ignorance and Gettier conditions of the Darrel vignette (Cramer’s = .02, 95% CI [.02, .06], (1) = 1.01, = .315), participants were less likely to attribute Darrel’s identification to luck in the knowledge condition (0.24) than in the Gettier condition (0.39; Cramer’s = .16, 95% CI [.13, .20], (1) = 77.40, < .001).

The likelihood of luck attributions also varied by condition matched with the Emma vignette, (2) = 706.98, p < .001. Participants were more likely to attribute Emma’s identification to luck in response to the Gettier condition (0.92) than in response to the ignorance condition (0.45; Cramer’s = .50, 95% CI [.47, .54], (1) = 713.55, < .001). Further, participants were less likely to attribute the outcome of the Emma vignette to luck in the knowledge condition (0.65) than in the Gettier condition (0.92; Cramer’s = .32, 95% CI [.29, .36], (1) = 298.19, < .001).

The likelihood of luck attributions also differed according to the Gerald vignette condition, (2) = 607.03, p < .001. Participants were more likely to attribute luck in response to the Gettier condition (0.92) than to the ignorance condition (0.45; Cramer’s = .20, 95% CI [.16, .23], (1) = 109.40, < .001). Participants were less likely to attribute Gerald’s identification to luck in the knowledge condition (0.65) than in the Gettier condition (Cramer’s = .20, 95% CI [.16, .23], (1) = 109.26, < .001). Thus, seemingly, the vignette by condition interaction was driven by responses to the Gettier condition: The difference in likelihoods of luck attributions between the Gettier and ignorance conditions was absent for the Darrel vignette, moderate for the Gerald vignette, and large for the Emma vignette.

bin\_cond <- stats::chisq.test(l.data\_graph$luck\_vas\_combined[l.data\_graph$luck\_vas\_combined\_source == "Binary"],  
 l.data\_graph$cond[l.data\_graph$luck\_vas\_combined\_source == "Binary"]) %>%   
 tidy()   
  
  
b\_table <- table(l.data\_graph$luck\_vas\_combined[l.data\_graph$luck\_vas\_combined\_source == "Binary"],  
 l.data\_graph$cond[l.data\_graph$luck\_vas\_combined\_source == "Binary"])  
  
bin\_cond\_GI <- prop.test(t(b\_table[1:2, 1:2])) %>%   
 tidy()  
  
bin\_cond\_GI\_v <- v.chi.sq(x2 = prop.test(t(b\_table[1:2, 1:2]))$statistic,   
 n = sum(t(b\_table[1:2, 1:2])),  
 r = 2, c = 2)   
  
bin\_cond\_GK<- prop.test(t(b\_table[1:2, c(1,3)])) %>%   
 tidy()  
  
bin\_cond\_GK\_v <- v.chi.sq(x2 = prop.test(t(b\_table[1:2, c(1,3)]))$statistic,   
 n = sum(t(b\_table[1:2, c(1,3)])),  
 r = 2, c = 2)

vas\_cond <- stats::chisq.test(l.data\_graph$luck\_vas\_combined[l.data\_graph$luck\_vas\_combined\_source == "VAS"],  
 l.data\_graph$cond[l.data\_graph$luck\_vas\_combined\_source == "VAS"]) %>%   
 tidy()   
  
  
vas\_table <- table(l.data\_graph$luck\_vas\_combined[l.data\_graph$luck\_vas\_combined\_source == "VAS"],  
 l.data\_graph$cond[l.data\_graph$luck\_vas\_combined\_source == "VAS"])  
  
vas\_cond\_GI <- prop.test(t(vas\_table[1:2, 1:2])) %>%   
 tidy()  
  
vas\_cond\_GI\_v <- v.chi.sq(x2 = prop.test(t(vas\_table[1:2, 1:2]))$statistic,   
 n = sum(t(vas\_table[1:2, 1:2])),  
 r = 2, c = 2)   
  
vas\_cond\_GK<- prop.test(t(vas\_table[1:2, c(1,3)])) %>%   
 tidy()  
  
vas\_cond\_GK\_v <- v.chi.sq(x2 = prop.test(t(vas\_table[1:2, c(1,3)]))$statistic,   
 n = sum(t(vas\_table[1:2, c(1,3)])),  
 r = 2, c = 2)

mutate(tidy(l.model.6C), model = "Model 6C") %>%   
 full\_join(  
 mutate(tidy(l.model.6B), model = "Model 6B"),   
 .) %>%   
 full\_join(  
 mutate(tidy(l.model.6A), model = "Model 6A"),  
 .) %>%   
 full\_join(  
 mutate(tidy(l.model.6), model = "Model 6"), .) %>%   
 full\_join(  
 mutate(tidy(l.model.5), model = "Model 5"), .) %>%   
 full\_join(  
 mutate(tidy(l.model.4), model = "Model 4"), .) %>%   
 full\_join(  
 mutate(tidy(l.model.3), model = "Model 3"), .) %>%   
 full\_join(  
 mutate(tidy(l.model.2), model = "Model 2"), .) %>%   
 full\_join(  
 mutate(tidy(l.model.1), model = "Model 1"), .) %>%   
 select(., model, effect, group, term, estimate, std.error, statistic, p.value) %>%   
 kable(caption = "Luck Attribution Model Results") %>%   
 kable\_styling()

## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("effect", "group", "term", "estimate", "std.error",  
## "statistic", "p.value", "model")  
## Joining, by = c("term", "estimate", "std.error", "statistic", "p.value",  
## "model")

Luck Attribution Model Results

model

effect

group

term

estimate

std.error

statistic

p.value

Model 1

NA

NA

(Intercept)

-0.0430442

0.0176638

-2.4368564

0.0148156

Model 2

fixed

NA

(Intercept)

-0.0442355

0.3359940

-0.1316556

0.8952567

Model 2

ran\_pars

vignette

sd\_\_(Intercept)

0.5810837

NA

NA

NA

Model 3

fixed

NA

(Intercept)

-0.0442355

0.3359968

-0.1316545

0.8952576

Model 3

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000002

NA

NA

NA

Model 3

ran\_pars

vignette

sd\_\_(Intercept)

0.5810886

NA

NA

NA

Model 4

fixed

NA

(Intercept)

-0.0442355

0.3359841

-0.1316595

0.8952536

Model 4

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0001126

NA

NA

NA

Model 4

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 4

ran\_pars

vignette

sd\_\_(Intercept)

0.5810667

NA

NA

NA

Model 5

fixed

NA

(Intercept)

0.4036010

0.3558809

1.1340900

0.2567568

Model 5

fixed

NA

compYes

-0.1777094

0.0428462

-4.1476116

0.0000336

Model 5

fixed

NA

age

0.0035660

0.0019174

1.8598358

0.0629088

Model 5

fixed

NA

gender2male

0.0074075

0.0411403

0.1800541

0.8571101

Model 5

fixed

NA

education

-0.0294521

0.0073157

-4.0258534

0.0000568

Model 5

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 5

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 5

ran\_pars

vignette

sd\_\_(Intercept)

0.5846120

NA

NA

NA

Model 6

fixed

NA

(Intercept)

-0.2463155

0.3722464

-0.6617002

0.5081634

Model 6

fixed

NA

compYes

-0.1806945

0.0439764

-4.1088941

0.0000398

Model 6

fixed

NA

age

0.0041529

0.0019673

2.1109937

0.0347729

Model 6

fixed

NA

gender2male

0.0072667

0.0421847

0.1722583

0.8632345

Model 6

fixed

NA

education

-0.0310962

0.0074846

-4.1546683

0.0000326

Model 6

fixed

NA

condIgnorance

1.0277337

0.0474143

21.6756170

0.0000000

Model 6

fixed

NA

condKnowledge

0.9356795

0.0471059

19.8633078

0.0000000

Model 6

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0991901

NA

NA

NA

Model 6

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0662757

NA

NA

NA

Model 6

ran\_pars

vignette

sd\_\_(Intercept)

0.6111192

NA

NA

NA

Model 6A

fixed

NA

(Intercept)

0.9073043

0.1285775

7.0564806

0.0000000

Model 6A

fixed

NA

compYes

-0.1953080

0.0446956

-4.3697376

0.0000124

Model 6A

fixed

NA

age

0.0041522

0.0020023

2.0737491

0.0381026

Model 6A

fixed

NA

gender2male

-0.0042108

0.0427246

-0.0985578

0.9214894

Model 6A

fixed

NA

education

-0.0310816

0.0076074

-4.0856980

0.0000439

Model 6A

fixed

NA

condIgnorance

0.0839290

0.0784178

1.0702794

0.2844936

Model 6A

fixed

NA

condKnowledge

0.7235533

0.0824618

8.7744067

0.0000000

Model 6A

fixed

NA

vignetteEmma

-2.8510151

0.1106195

-25.7731706

0.0000000

Model 6A

fixed

NA

vignetteGerald

-1.0704438

0.0788193

-13.5809879

0.0000000

Model 6A

fixed

NA

condIgnorance:vignetteEmma

2.5277977

0.1350603

18.7160679

0.0000000

Model 6A

fixed

NA

condKnowledge:vignetteEmma

1.0673419

0.1382340

7.7212696

0.0000000

Model 6A

fixed

NA

condIgnorance:vignetteGerald

0.7232640

0.1105027

6.5452171

0.0000000

Model 6A

fixed

NA

condKnowledge:vignetteGerald

0.0949037

0.1137941

0.8339948

0.4042839

Model 6A

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0215039

NA

NA

NA

Model 6A

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0395744

NA

NA

NA

Model 6A

ran\_pars

vignette

sd\_\_(Intercept)

0.0000000

NA

NA

NA

Model 6B

fixed

NA

(Intercept)

-0.3697953

0.3730208

-0.9913529

0.3215133

Model 6B

fixed

NA

compYes

-0.1215943

0.0457124

-2.6599846

0.0078144

Model 6B

fixed

NA

age

0.0086803

0.0021946

3.9553613

0.0000764

Model 6B

fixed

NA

gender2male

0.0335005

0.0426260

0.7859178

0.4319156

Model 6B

fixed

NA

education

-0.0320762

0.0074960

-4.2791019

0.0000188

Model 6B

fixed

NA

condIgnorance

1.0474525

0.0491591

21.3074051

0.0000000

Model 6B

fixed

NA

condKnowledge

0.9230960

0.0488703

18.8887016

0.0000000

Model 6B

fixed

NA

turkTRUE

-0.3788377

0.1411397

-2.6841322

0.0072718

Model 6B

fixed

NA

condIgnorance:turkTRUE

-0.2791931

0.1877634

-1.4869415

0.1370303

Model 6B

fixed

NA

condKnowledge:turkTRUE

0.1895728

0.1831578

1.0350245

0.3006575

Model 6B

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0730888

NA

NA

NA

Model 6B

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0898205

NA

NA

NA

Model 6B

ran\_pars

vignette

sd\_\_(Intercept)

0.6106858

NA

NA

NA

Model 6C

fixed

NA

(Intercept)

-0.1509830

0.3816524

-0.3956034

0.6923976

Model 6C

fixed

NA

compYes

-0.1799466

0.0440499

-4.0850640

0.0000441

Model 6C

fixed

NA

age

0.0043256

0.0019697

2.1961382

0.0280821

Model 6C

fixed

NA

gender2male

0.0058848

0.0422613

0.1392481

0.8892541

Model 6C

fixed

NA

education

-0.0300051

0.0074939

-4.0039265

0.0000623

Model 6C

fixed

NA

condIgnorance

1.4199932

0.1266963

11.2078518

0.0000000

Model 6C

fixed

NA

condKnowledge

0.8159375

0.1233020

6.6173892

0.0000000

Model 6C

fixed

NA

luck\_vas\_combined\_sourceVAS

-0.1349282

0.0954996

-1.4128670

0.1576948

Model 6C

fixed

NA

condIgnorance:luck\_vas\_combined\_sourceVAS

-0.4538761

0.1363798

-3.3280310

0.0008746

Model 6C

fixed

NA

condKnowledge:luck\_vas\_combined\_sourceVAS

0.1406571

0.1332710

1.0554219

0.2912324

Model 6C

ran\_pars

person\_code:(id:vignette)

sd\_\_(Intercept)

0.0724694

NA

NA

NA

Model 6C

ran\_pars

id:vignette

sd\_\_(Intercept)

0.0980260

NA

NA

NA

Model 6C

ran\_pars

vignette

sd\_\_(Intercept)

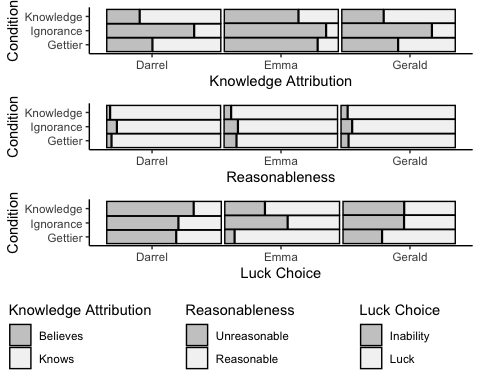
0.6132810

NA

NA

NA

cowplot::plot\_grid(k.interact + theme(legend.position="none") +   
 xlab("Knowledge Attribution"),  
 r.interact + theme(legend.position="none") +   
 xlab("Reasonableness"),  
 l.interact + theme(legend.position="none") +   
 xlab("Luck Choice"),   
 plot\_grid(get\_legend(k.interact),  
 get\_legend(r.interact),  
 get\_legend(l.interact),  
 nrow = 1),  
 nrow = 4)



**Luck as a Moderator.** Next, we explored whether attributions of outcomes to luck versus ability influence knowledge attributions as suggested by prior research (Turri, 2016, 2017). Turri (2016; Experiment 7) found a strong positive correlation between knowledge attributions and attributions to ability rather than luck (r = .622) and a moderating effect of luck attributions on Gettier intuitions; the participants attributed knowledge less to protagonists that were perceived as having arrived at a truth because of a lucky guess rather than because of their ability (ηp2 = .353; Turri, 2016; Experiment 7).

# luck  
luck\_interaction\_cond <- stats::chisq.test(l2.data\_graph$know\_vas\_combined[l2.data\_graph$luck\_vas\_combined == "Luck"],   
 l2.data\_graph$cond[l2.data\_graph$luck\_vas\_combined == "Luck"]) %>%   
 tidy()  
  
  
luck\_interaction\_table <- table(l2.data\_graph$know\_vas\_combined[l2.data\_graph$luck\_vas\_combined == "Luck"], l2.data\_graph$cond[l2.data\_graph$luck\_vas\_combined == "Luck"])  
  
  
luck\_interaction\_cond\_GI <- prop.test(t(luck\_interaction\_table[2:1, 1:2])) %>% tidy()  
  
luck\_interaction\_cond\_GI\_v <- v.chi.sq(x2 = prop.test(t(luck\_interaction\_table[2:1, 1:2]))$statistic,   
 n = sum(t(luck\_interaction\_table[2:1, 1:2])),  
 r = 2, c = 2)  
  
  
luck\_interaction\_cond\_GK <- prop.test(t(luck\_interaction\_table[2:1, c(1,3)])) %>% tidy()  
  
luck\_interaction\_cond\_GK\_v <- v.chi.sq(x2 = prop.test(t(luck\_interaction\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(luck\_interaction\_table[2:1, c(1,3)])),  
 r = 2, c = 2)  
  
#ability  
ability\_interaction\_cond <- stats::chisq.test(l2.data\_graph$know\_vas\_combined[l2.data\_graph$luck\_vas\_combined == "Ability"],   
 l2.data\_graph$cond[l2.data\_graph$luck\_vas\_combined == "Ability"]) %>%   
 tidy()  
  
  
ability\_interaction\_table <- table(l2.data\_graph$know\_vas\_combined[l2.data\_graph$luck\_vas\_combined == "Ability"], l2.data\_graph$cond[l2.data\_graph$luck\_vas\_combined == "Ability"])  
  
  
ability\_interaction\_cond\_GI <- prop.test(t(ability\_interaction\_table[2:1, 1:2])) %>% tidy()  
  
ability\_interaction\_cond\_GI\_v <- v.chi.sq(x2 = prop.test(t(ability\_interaction\_table[2:1, 1:2]))$statistic,   
 n = sum(t(ability\_interaction\_table[2:1, 1:2])),  
 r = 2, c = 2)  
  
  
ability\_interaction\_cond\_GK <- prop.test(t(ability\_interaction\_table [2:1, c(1,3)])) %>% tidy()  
  
ability\_interaction\_cond\_GK\_v <- v.chi.sq(x2 = prop.test(t(ability\_interaction\_table[2:1, c(1,3)]))$statistic,   
 n = sum(t(ability\_interaction\_table[2:1, c(1,3)])),  
 r = 2, c = 2)

We tested whether luck attributions moderated the effect of condition on knowledge attribution among participants who correctly identified that the protagonist was right in their identification. The main effect of luck attributions and the interaction between condition and luck attributions were added to Model 6 of the knowledge attributions analysis (Model 6D; AIC = 13363.98). This model (Pseudo-*R*2 =0.24 - 0.28) explained more variance in knowledge attributions than Model 6.

Condition affected knowledge attributions when participants attributed the protagonists’ correct identification to luck, (2) = 211.03, p < .001. Participants were more likely to attribute knowledge to the protagonist in the Gettier condition vignette (0.31) than in the ignorance condition vignette (0.17; Cramer’s = .16, 95% CI [.13, .19], (1) = 116.03, < .001). They were also more likely to attribute knowledge in the knowledge condition vignette (0.38) than in the Gettier condition vignette (0.31; Cramer’s = .07, 95% CI [.05, .10], (1) = 24.54, < .001).

Similarly, condition affected knowledge attributions when participants attributed the protagonists’ correct identification to ability (2) = 1737.19, p < .001. Participants were more likely to attribute knowledge to the protagonist in the Gettier condition vignette (0.66) than in the ignorance condition vignette (0.16; Cramer’s = .51, 95% CI [.48, .54], (1) = 972.07, < .001). Participants were also more likely to attribute knowledge in the knowledge condition vignette (0.73) than in the Gettier condition vignette (0.66; Cramer’s = .08, 95% CI [.05, .11], (1) = 21.36, < .001). Both of these effects were larger than those found when participants attributed the outcome to luck. As found in previous research, luck attributions attenuated the effect of condition on knowledge.

*Alternative Knowledge Probe*

Next, we assessed whether question wording affected participants’ knowledge attributions as has been suggested by previous research (e.g., Machery et al., 2017; Nagel, San Juan, et al., 2013). Participants may be more likely to deny knowledge to a protagonist when they are asked a more nuanced question (whether the protagonist knew or only felt like they knew but did not actually know; Nagel, San Juan, et al., 2013) than when they are asked a simpler question (whether the protagonist knew or did not know).

In our exploratory analyses of the alternative knowledge probe (i.e., following Model steps 1 through 6 for the other analyses, we found a pattern of results similar to those for the analyses of our primary knowledge measure (Model 6D: AIC = 16332.68; Pseudo-*R*2 =0.16 - 0.21). In response to the Gettier condition vignettes and the ignorance condition vignette, participants were less likely to indicate that the protagonist had knowledge than to indicate that the protagonist felt they had knowledge but did not have it. In the knowledge condition vignettes, people were fairly equal in responding that the protagonist knew and stating that the protagonist felt they knew, but didn’t know.

table(k.model.6F@frame$cond, k.model.6F@frame$know\_alt) %>%   
 kable(caption = "Alternative Knowledge Probe", col.names = c("Knows", "Feels they know, but doesn't know")) %>%   
 kable\_styling()

Alternative Knowledge Probe

Knows

Feels they know, but doesn’t know

Gettier

1871

2879

Ignorance

840

3910

Knowledge

2310

2441

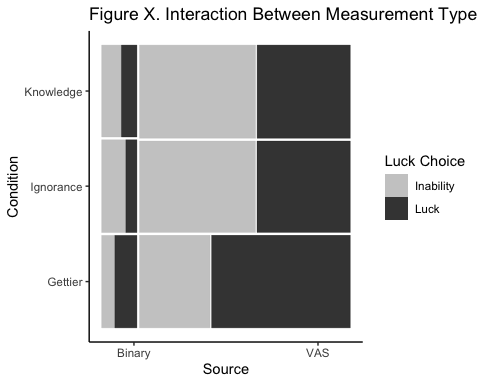
*Measurement Characteristics*

Finally, we examined whether condition effects were influenced by measurement characteristics, specifically if the outcome was originally measured on a binary or visual analogue scale. Adding measurement and its interaction with condition to the model predicting knowledge attribution did not improve model fit (Model 6C; AIC = 15876.57); thus, we found no evidence for moderation (Pseudo-*R*2 =0.21 - 0.25). Next, we estimated the same model (Model 6C) in predicting judgments of reasonableness (AIC = 7041.29). While this model performed better than Model 6, the interactions between condition and measurement type were not significant (Pseudo-*R*2 =0.02 - 0.07). Finally, we estimated a model that included an interaction between condition and measurement type predicting luck attributions (Model 6C, AIC = 15862.09). This model performed better than Model 6 (Pseudo-*R*2 =0.14 - 0.16) and revealed an interaction effect for the Ignorance condition in comparison to the Gettier condition. As shown in Figure X, the effect of condition was consistent in direction across measurement type but appears to have differed in magnitude.

Condition affected the likelihood of luck attributions on responses to the binary measure, (2) = 120.98, p < .001. Participants were more likely to attribute outcomes to luck in the Gettier condition (0.37) than in the ignorance condition (0.68; Cramer’s = .31, 95% CI [.26, .37], (1) = 118.14, < .001). Participants were also more likely to attribute outcomes to luck in the Gettier condition (0.37) than in the knowledge condition (0.55; Cramer’s = .18, 95% CI [.13, .24], (1) = 41.00, < .001).

Condition similarly affected luck attributions as measured by the VAS, (2) = 454.78, p < .001. Participants were more likely to attribute outcomes to luck in the Gettier condition (0.34) than in the ignorance condition (0.55; Cramer’s = .22, 95% CI [.20, .24], (1) = 341.27, < .001) Participants were also more likely to attribute outcomes to luck in the Gettier condition (0.34) than in the knowledge condition (0.55; Cramer’s = .22, 95% CI [.20, .24], (1) = 345.90, < .001). Effect sizes were smaller when luck was measured continuously, but the confidence intervals of these effect sizes overlapped with those produced by the binary measure.

#graph the three way binary   
ggplot(l.data\_graph) +  
 geom\_mosaic(aes(x = product(luck\_vas\_combined, cond, luck\_vas\_combined\_source),   
 fill = luck\_vas\_combined)) +   
 scale\_fill\_manual(name = "Luck Choice",  
 values = c("gray", "black")) +   
 scale\_x\_productlist(breaks = c(0.13,.87),  
 labels = c("Binary", "VAS")) +   
 theme\_classic() +   
 xlab("Source") +   
 ylab("Condition") +  
 ggtitle("Figure X. Interaction Between Measurement Type and Condition on Luck")



# Disclosures

*Data, materials, and online resources:* Deidentified raw data and deidentified data with exclusions are are posted publicly on our master OSF page (<https://osf.io/n5b3w/>), and each contributing site has posted their data on an OSF page linked to our master OSF page.

*Reporting:*

“We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study” (see Simmons, Nelson, & Simonsohn, 2011).

*Ethical approval:*

All contributing labs were required to submit their local institutional ethics approval prior to data collection as part of their pre-registration and CREP review process and will be carried out in accordance with the provisions of the World Medical Association Declaration of Helsinki. All participating labs posted their ethics approval to their lab’s OSF page for this study.

*Author Contributions:*

TBD

*Conflicts of Interest:*

The author(s) declare that there were no conflicts of interest with respect to the authorship or the publication of this article.

*Funding:*

TBD

*Supplemental Material:*

If Supplemental Material will be posted on the journal’s Web site, include this heading and the appropriate link will be added during editing.

*Prior versions:*

If part or all of a submitted manuscript was previously posted to a blog or to a preprint archive, the authors should provide a link to that source and briefly indicate what aspects of the submitted manuscript are shared with that prior version.

# Appendix

### Vignettes

*For the Gerald vignette, all participants read:*

“Gerald is driving through the countryside with his young son Andrew. Along the way he sees numerous objects and points them out to his son. ‘That’s a cow, Andrew,’ Gerald says, ‘and that over there is a house where farmers live.’ Gerald has no doubt about what the objects are.

What Gerald and Andrew do not realize is the area they are driving through was recently hit by a very serious tornado. This tornado did not harm any of the animals, but did destroy most buildings. In an effort to maintain the rural area’s tourist industry, local townspeople built new houses in the place of the destroyed houses. These new houses were rebuilt with all the materials necessary for them to look exactly like the original houses from the road, and they are also fully furnished and can now be used as actual housing.”

In the knowledge condition, participants read:

“Having just entered the tornado-ravaged area, Gerald notices the many houses lining the roads. When he tells Andrew ‘That’s a house,’ the object he sees and points at is a real house that has survived the tornado and not one of the new houses.”

In the ignorance condition, participants read:

“Having driven through the tornado-ravaged area, Gerald has encountered many of these fake houses. When he tells Andrew ‘That’s a house,’ the object he sees and points at is a fake house that was built after the tornado and is not actually a house.”

In the Gettier condition, participants read:

“Having just entered the tornado-ravaged area, Gerald has not yet encountered any fake houses. When he tells Andrew ‘That’s a house,’ the object he sees and points at is a real house that has survived the tornado and not one of the fake houses.”

*For the Emma vignette, all participants read:*

“Emma is shopping for jewelry. She goes into a nice-looking store and selects a necklace from a tray marked”Diamond Earrings and Pendants.” “What a lovely diamond!” she says as she tries it on. Emma could not tell the difference between a real diamond and a cubic zirconium fake just by looking or touching.”

In the knowledge condition, participants read:

“However, this particular store has very honest employees who have a really positive reputation for their guaranteed real diamonds; in the tray Emma chose, all of the pendants had real diamonds rather than fake cubic zirconium stones (and the one she chose was really nice).”

In the ignorance condition, participants read:

“Unfortunately, this particular store has very dishonest employees who have been stealing real diamonds and replacing them with fakes; in the tray Emma chose, almost all of the pendants had cubic zirconium stones rather than diamonds (and the one she chose was in fact fake).”

In the Gettier condition, participants read:

“Unfortunately, this particular store has very dishonest employees who have been stealing real diamonds and replacing them with fakes; in the tray Emma chose, almost all of the pendants had cubic zirconium stones rather than diamonds (but the one she chose happened to be real).”

*For the Darrel vignette, all participants read:*

“Darrel is an ecologist collecting data on red speckled ground squirrels in Canyon Falls national park. The park is divided into ten zones and today Darrel is working Zone 3. While scanning the river valley with his binoculars, Darrel sees a small, bushy-tailed creature with distinctive red markings on its chest and belly. The red speckled ground squirrel is the only native species with such markings. Darrel records in his journal, ‘At least one red speckled ground squirrel in Zone 3 today.’

In the knowledge condition, participants read:

“Ecologists are unaware that a complex network of aquifers recently began drying up in the park. These aquifers carry vital nutrients to the trees and other forms of plant life that support the squirrels. And the aquifers in the river valley running through Zone 3 are no exception. The animal Darrel is looking at is indeed a thirsty red speckled ground squirrel.”

In the ignorance condition, participants read:

“Ecologists are unaware that a non-native species of prairie dog recently began invading the park. These prairie dogs also have red markings on their chest and belly. When these prairie dogs tried to invade Zone 3, the red speckled ground squirrels were unable to completely drive them away. And, the animal Darrel is looking at is indeed one of the prairie dogs.”

In the Gettier condition, participants read:

“Ecologists are unaware that a non-native species of prairie dog recently began invading the park. These prairie dogs also have red markings on their chest and belly. When these prairie dogs tried to invade Zone 3, the red speckled ground squirrels were unable to completely drive them away. Still, the animal Darrel is looking at is a red speckled ground squirrel.”

### Exclusion ratings

**How exclusions were made**

No or NA gets 0 points Maybe gets 1 point Yes/test gets 2 points Participants are marked as “excluded” if they get 4 total points across three coders

Participants are also marked as “nonsense” if they did not write a legible answer or simply typed gibberish. These data points are not excluded but marked.

**Instructions Given to Raters**

Note: These instructions were adapted from instructions written by William McAuliffe and Hannah Moshontz for an unrelated project

We need help coding open-ended responses that will inform our pre-registered exclusion criteria for this project. Specifically, we will exclude data on the basis of a suspicion check (whether people guess the study hypothesis) and previous study participation (whether people describe having participated in similar studies before).

All participants were asked two questions (with some labs asking slight variations): What is, in your opinion, the purpose of this study? (purpose) Have you ever participated in a similar study? If yes, please describe the study. (previous)

Your task is to evaluate people’s answers to these questions. We will have 2 people evaluate every response and then we will exclude people based on the average.

For each question this is how we would like you to evaluate answers. If you are coding from a language other than English, please directly assess the question (rather than translating it) and provide a code/label in English (yes, maybe, no, test, as described below).

**purpose**

*yes*

The participant identified that we are studying justified true belief and Gettier cases.

Example “yes” coding cases: “To test exceptions to the Justified True Belief theory”

*maybe*

The participant describes something similar to the true study hypothesis (true knowledge is different from a lucky or incidentally correct belief).

Example “maybe” coding cases: “To see if a story can change ones perception of knowledge based on luck or ability”

*no*

The participant did not identify the study hypothesis or offer a very vague description, which might include the words belief or knowledge.

Example “no” coding cases: “I think the purpose was to see how do people classify if someone knows something or if they just strongly agree with it”; OR “understand how people view scenarios based on the words used to describe them”

*test*

The response indicates that it is a test case

Example “test” coding case: “TEST”; OR “test”; OR “this is a test”

*NA*

If you are unsure how to code a response, you can write NA.

Example “NA” coding case: “nnnnnnnnnnn”

**Previous**

*yes*

The participant has participated in this exact study before, or an exact replication of it.

Example “yes” coding cases: “Yes, I completed this study before.”

*maybe*

The participant has participated in a similar study before, or may have based on their description.

Example “maybe” coding cases: “Yes another study that was very similar.”; “Yes, I have participated in a study that asked similar questions but had slightly different scenarios”

*no*

The participant has not participated in this study or a similar study before based on their description.

Example “no” coding cases: “nope”; “Yes, I have participated in a study for course credit before.”’; “Yes, I have done studies where I read scenarios and answered questions about them.”

*test*

The response indicates that it is a test case

Example “test” coding case: “TEST”; OR “test”; OR “this is a test”

*NA*

If you are unsure how to code a response, you can write NA.

Example “NA” coding case: “nnnnnnnnnnn”

**Do’s and Don’ts** Take breaks! This work is hopefully interesting, but it can be cognitively exhausting. If you are having trouble paying attention while you are doing this or if you feel tired of it, please take a break.

After you label a response, do not go back and change it later. This may be tempting to do after mentally comparing how you rated different responses, but just carefully work through each response and know that your initial rating is final.

Don’t discuss your ratings with other raters—this will invalidate everyone’s work. Do assign labels for every response in your assigned sheet(s). If you would like to contribute more, please email the person\_code listed at the top of this sheet.

**To summarize, for each set of answers** Read the answer to the question and assign a label that describes either whether people intuited the study hypothesis (for purpose) or whether people participated in a similar study previously (for previous)

**Coding form includes** [id] A subject id number [survey\_lang] [purpose/previous] The answer people gave to the question [code] The code/label that you are assigning to the answer (yes, no, maybe, test)

### Demographics

*Version A* Used by the following data collection sites in SSS: [List sites that used this option here]

White / European descent Black / African descent Latino\*a / Latin American descent Australian descent Asian Southeast Asian descent Native American Hawaiian descent / Pacific Islands Other

*Version B* Used by the following data collection sites in SSS: [List sites that used this option here]

European descent African descent Latino\*a / Latin American descent Indigenous Australian or Torres Strait Islander descent East Asian descent South Asian descent Pacific Island descent Native American descent

*Version C* Used by the following data collection sites in Qualtrics: [List sites that used this option here] White/European Black/African American Hispanic Latino East or Southeast Asian / Pacific Islander (e.g. from Japan, China, Korea, Vietnam, Thailand, Philippines, native Hawaiian) South Asian (e.g. from India, Pakistan) I prefer not to answer this question Other

1. The Stage 1 registered report manuscript included sections that described the recruitment and approval of collaborators who would collect data. We have restructured the Method section to more closely resemble that of a typical empirical article. The original text, updated to reflect the study’s completion, can be found in Appendix A. [↑](#footnote-ref-90)
2. In the approved protocol, we described a plan for data collection whereby each lab pre-registered a target sample size of 50-100 and stopped collecting data on April 1, 2020 or once all contributors reached their pre-registered target sample size. Due to the COVID-19 pandemic, this plan was not followed. The deadline for data collection was extended to June 1, 2021. Many data collection sites stopped collecting data earlier. [↑](#footnote-ref-91)
3. There may be measurement error in participants reported years of education. 393 people report what amounts to less than a high school diploma. 66 people said they had one year of education. [↑](#footnote-ref-92)
4. While we planned to measure participants’ racial and ethnic identities using an open-ended response, racial and ethnic identity was measured using non-exclusive categories with an open-ended fill-in option for reasons that were not documented. Student research teams designed different response options tailored to their geographic region (see all variations in Appendix Table X). All data collection sites allowed people to select multiple racial and ethnic identities, and all asked whether participants identified as White (either “White/European”, “White / European descent”, or “European descent”). [↑](#footnote-ref-93)
5. We did not pre-register the exclusion of people who reported their age as over 100; 7 people were excluded on the basis of this criteria alone (i.e., they did not meet any other exclusion criteria). These responses may have been errors in data entry, or they may have been unlabeled test responses. [↑](#footnote-ref-94)
6. Data collection sites were not given instructions about avoiding or clearly identifying test responses. At many data collection sites, the students and other researchers executing the study tested their survey link multiple times (e.g., as inferred by responses to open-ended questions marked “test”). [↑](#footnote-ref-96)
7. The approved protocol described a power analysis conducted prior to data collection. The text from the original protocol is reproduced in full in Appendix X, and is summarized here. [↑](#footnote-ref-97)
8. Teams that participated in this extension were required to collect twice as many participants so that they could meet the sample size requirement (N = 50) for participants using only the pre-approved continuous measure (half in the continuous condition and half in the binary condition; N > 100). However, because we converted all continuous responses to binary responses (see Analytic Approach section below for more details), the binary responses collected using this extension were also included alongside the converted binary responses. [↑](#footnote-ref-98)
9. For participants from the United States, less than a high- school education was coded as 10 years, a high school diploma was coded as 12 years, some college or a 2-year college degree was coded as 14 years, a 4-year college degree was coded as 16 years, a master’s was coded as 18 years, and a doctorate or professional degree was coded as 20 years. For participants from Portugal, the labels and coding were the same except that a 4-year college degree was coded as 15 years and a doctorate degree was coded to 21 years. [↑](#footnote-ref-99)
10. In the approved protocol, data collection was described as taking place in labs. Labs were described as uniquely identifying data collection sites. However, at some data collection sites, multiple student-led teams joined this project (e.g., under the mentorship of the same PI, multiple students joined the project as “labs”). Observations were labeled as belonging to both a “lab” (which we describe as a “student-led team”) and a data collection site. For analyses, the data collection site was used in place of the “lab” variable described in the approved protocol. [↑](#footnote-ref-101)
11. In the approved protocol, the results section focused heavily on the project’s logistics and structured results reporting in ways that would not allow for a transparent and thorough description of model fit and other important aspects of results, like assumption checks. Further, some model specification details in the approved protocol conflicted with stated research questions (e.g., we specified that the null model would include the focal predictor - which would have rendered the null model invalid - as null models are not supposed to include any predictor). [↑](#footnote-ref-106)