R Figures

Contents

Import the data

```
# Import the longtable with all results
lt <- read.csv(file="total_longtable.csv")
# Separate out results for RQ2 and RQ3
lt_RQ2 <- lt[lt$initial_visualizations == 1,]
lt_RQ3 <- lt[lt$initial_visualizations == 0,]
```

Aggregating scores

We aggregate the longtable scores for comprehension, trust, and operationalized bias across all questions per participant to get their totals for every person.

```
lt_RQ2_aggregate <- lt_RQ2 %>%
  group_by(Pid, Vis_Type, qualitative_model_perception, Gender, Age,
           Familiarity) %>%
  summarise(
    comprehension = sum(comprehension),
    trust = sum(trust),
    likert.only.trust=sum(likert.only.trust),
    bias.operational = sum(bias.operational),
    correct.output = sum(correct.output),
    correct.pushing = sum(correct.pushing),
    correct.power = sum(correct.power),
    .groups = "drop"
  )
```

Mediation model

We fit a mediation model with trust score as the dependent variable, the comprehension score as the predictor, and the bias perception score as the mediator. We find that comprehension has a direct positive effect on bias perception, and bias perception had a direct negative effect on trust. That is, increased comprehension indirectly decreases trust by impacting perception of bias.

```
mediation_model <- '
  # Mediator equation: effect of comprehension on the mediator (bias percep.)
  bias.operational ~ a * comprehension
```

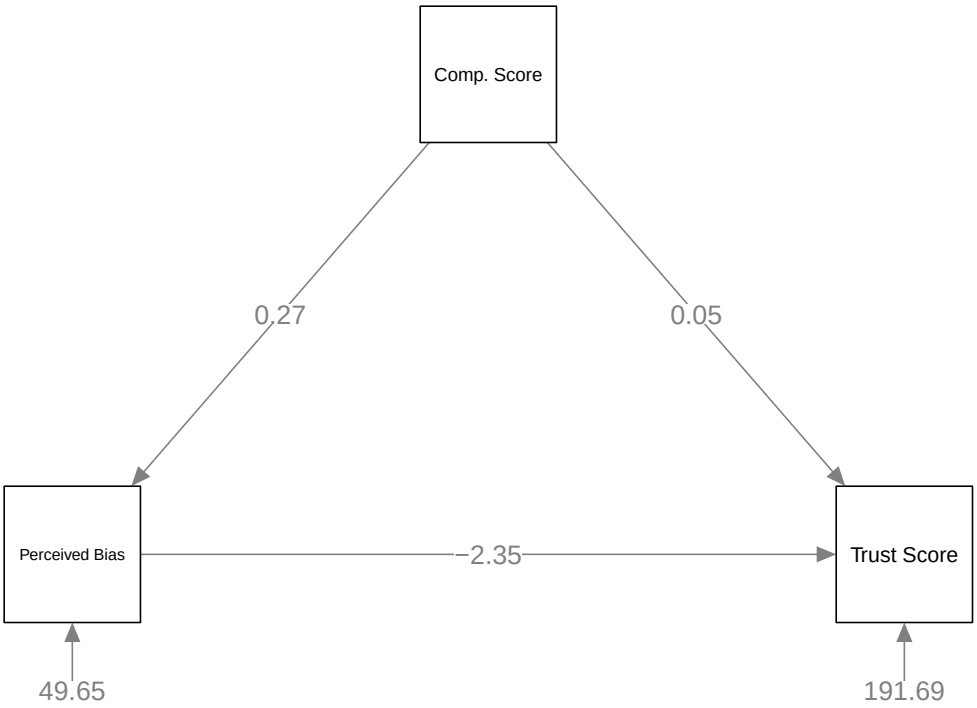
```
  # Outcome equation: direct effect of comprehension and effect of the
  # mediator (bias percep.) on trust
  trust ~ c_prime * comprehension + b * bias.operational

  # Indirect effect: the mediation path (a * b)
  indirect := a * b

  # Total effect: sum of the direct and indirect effects
  total := c_prime + indirect
'

# Estimate the mediation model
mediation_results <- sem(mediation_model, data = lt_RQ2_aggregate)

# Summarize the results
summary(mediation_results, standardized = TRUE, fit.measures = TRUE)
```

```
## lavaan 0.6-19 ended normally after 1 iteration
##
##   Estimator                                         ML
##   Optimization method                           NLMINB
##   Number of model parameters                         5
##
##   Number of observations                           444
##
## Model Test User Model:
##
##   Test statistic                                 0.000
##   Degrees of freedom                                 0
##
## Model Test Baseline Model:
##
##   Test statistic                               497.251
##   Degrees of freedom                                 3
##   P-value                                        0.000
##
## User Model versus Baseline Model:
##
##   Comparative Fit Index (CFI)                    1.000
##   Tucker-Lewis Index (TLI)                       1.000
##
## Loglikelihood and Information Criteria:
##
##   Loglikelihood user model (H0)              -3293.732
##   Loglikelihood unrestricted model (H1)      -3293.732
##
##   Akaike (AIC)                                6597.465
##   Bayesian (BIC)                              6617.944
##   Sample-size adjusted Bayesian (SABIC)       6602.076
##
## Root Mean Square Error of Approximation:
##
##   RMSEA                                          0.000
```

```
##   90 Percent confidence interval - lower         0.000
##   90 Percent confidence interval - upper         0.000
##   P-value H_0: RMSEA <= 0.050                       NA
##   P-value H_0: RMSEA >= 0.080                       NA
##
## Standardized Root Mean Square Residual:
##
##   SRMR                                           0.000
##
## Parameter Estimates:
##
##   Standard errors                             Standard
##   Information                                 Expected
##   Information saturated (h1) model          Structured
##
## Regressions:
##                     Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##   bias.operational ~
##     cmprhns    (a)      0.267    0.031    8.488    0.000    0.267    0.374
##   trust ~
##     cmprhns (c_pr)      0.046    0.067    0.690    0.490    0.046    0.022
##     bs.prtn    (b)     -2.355    0.093  -25.255    0.000   -2.355   -0.796
##
## Variances:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##    .bias.operatinl   49.650    3.332   14.900    0.000   49.650    0.860
##    .trust           191.689   12.865   14.900    0.000  191.689    0.379
##
## Defined Parameters:
##                    Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
##     indirect         -0.628    0.078   -8.046    0.000   -0.628   -0.297
##     total            -0.582    0.096   -6.041    0.000   -0.582   -0.276
```

We plot this mediation model below.

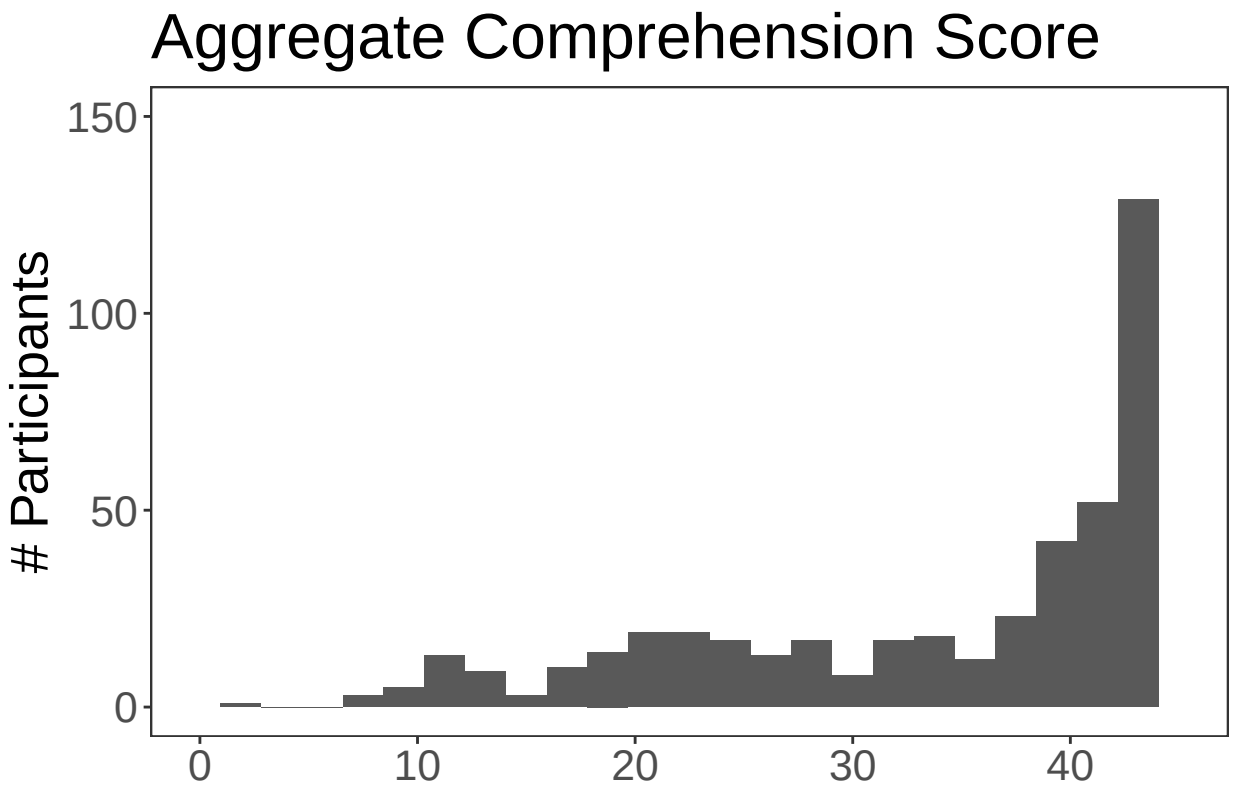
```
plot = semPaths(mediation_results, whatLabels = "est", style = "lisrel",
                nCharNodes=0,sizeMan=12, label.cex=1, sizeInt = 7,
                edge.label.cex=1, posCol='green', negCol='red',
                nodeLabels=c("Perceived Bias", "Trust Score", "Comp. Score"))
```



We also plot the distributions of each measure

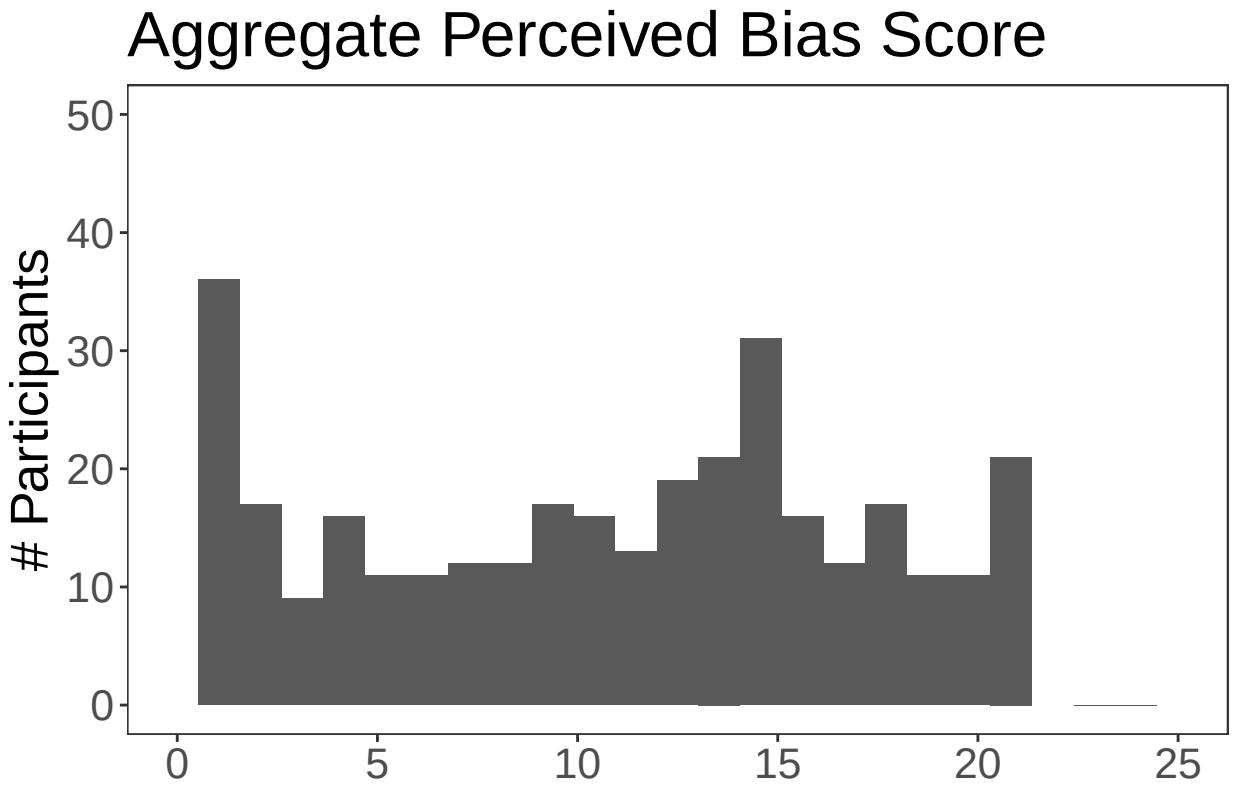
```
ggplot(lt_RQ2_aggregate, aes(comprehension)) +
  ggtitle("Aggregate Comprehension Score") +
  geom_histogram(bins=25) +
  ylim(0, 150) +
  xlim(0, 45) +
  theme_bw() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  theme(text = element_text(size = 20))+
  labs(x = "", y="# Participants")
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_bar()`).
```




```
ggplot(lt_RQ2_aggregate, aes(bias.operational)) +
  ggtitle("Aggregate Perceived Bias Score") +
  geom_histogram(bins=25) +
  ylim(0, 50) +
  xlim(0, 25) +
  theme_bw() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  theme(text = element_text(size = 20))+
  labs(x = "", y="# Participants")
```

```
## Warning: Removed 3 rows containing missing values or values outside the scale range
## (`geom_bar()`).
```




```
ggplot(lt_RQ2_aggregate, aes(trust)) +
  ggtitle("Aggregate Trust Score") +
  geom_histogram(bins=25) +
  ylim(0, 50) +
  xlim(0, 105)+
  theme_bw() +
  theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) +
  theme(text = element_text(size = 20))+
  labs(x = "", y="# Participants")
```

```
## Warning: Removed 2 rows containing missing values or values outside the scale range
## (`geom_bar()`).
```

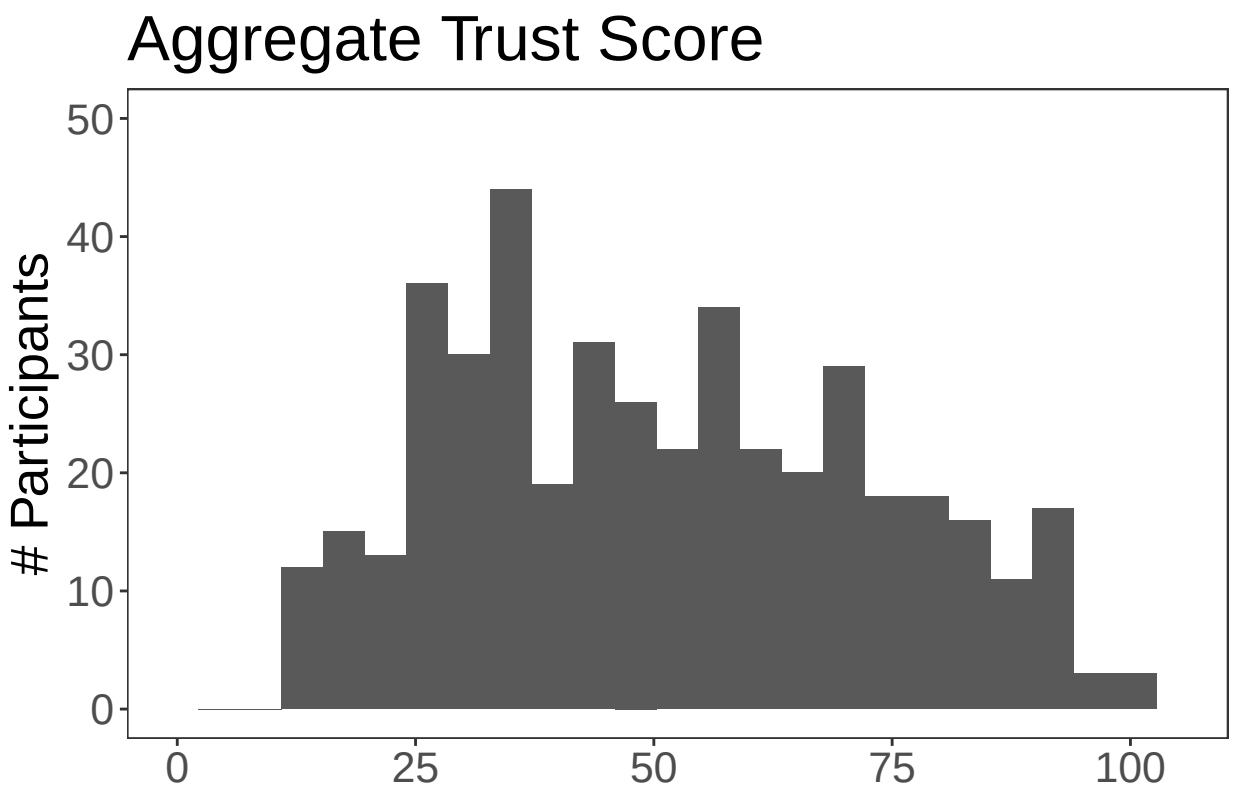


Figure 1

We plot the comprehension, bias perception and trust measures to be able to create figure 1

Setup

```
options(repr.plot.height=20)
lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(Vis_Type = case_when((Vis_Type == "shap") ~ "SHAP Waterfall", .default = Vis_Type))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(Vis_Type = case_when((Vis_Type == 'lime') ~ "LIME", .default = Vis_Type))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(Vis_Type = case_when((Vis_Type == 'forceshap') ~ "SHAP Force", .default = Vis_Type))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(Vis_Type = case_when((Vis_Type == 'eli5') ~ "ELI5", .default = Vis_Type))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(Vis_Type = case_when((Vis_Type == 'cp') ~ "Ceteris Paribus", .default = Vis_Type))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(Vis_Type = case_when((Vis_Type == 'anchors') ~ "Anchors", .default = Vis_Type))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(leastmost = case_when((Vis_Type == "SHAP Waterfall") ~ 3, .default = 0))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
```

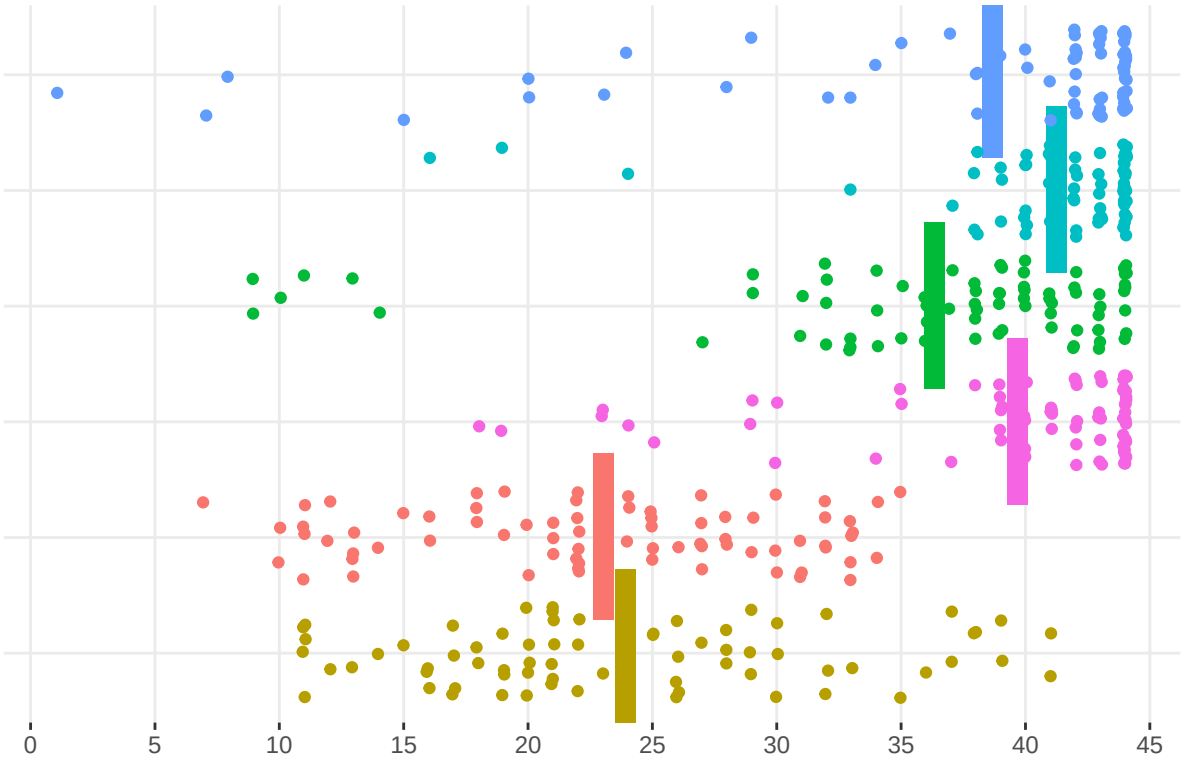
```
  mutate(leastmost = case_when((Vis_Type == 'LIME') ~ 5, .default = leastmost))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(leastmost = case_when((Vis_Type == 'SHAP Force') ~ 6, .default = leastmost))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(leastmost = case_when((Vis_Type == 'ELI5') ~ 4, .default = leastmost))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(leastmost = case_when((Vis_Type == 'Ceteris Paribus') ~ 1, .default = leastmost))

lt_RQ2_aggregate <- lt_RQ2_aggregate %>%
  mutate(leastmost = case_when((Vis_Type == 'Anchors') ~ 2, .default = leastmost))
```

Comprehension

```
lt_RQ2_aggregate %>% ggplot(aes(x = comprehension, y = reorder(Vis_Type, leastmost), color=Vis_Type)) +
  stat_summary(geom = "point", shape=108, fun.x = "mean", size = 30) +
  geom_jitter(width=0.08) +
  labs(
    title = "",
    x = "",
    y = ""
  ) +
  scale_y_discrete(name ="") +
  theme_bw() +
  theme(legend.key = element_blank(), strip.background = element_rect(colour="white", fill="white")) +
  theme(strip.text.y.left = element_text(angle = 0) ) +
  theme(legend.position = "none") +
  labs(x = "") +
  scale_x_continuous(breaks = seq(0, 45, by = 5)) +
  theme(panel.grid.minor = element_blank(), panel.border = element_blank(), axis.text.y=element_blank()
```

```
## Warning in stat_summary(geom = "point", shape = 108, fun.x = "mean", size =
## 30): Ignoring unknown parameters: `fun.x`
```

```
## No summary function supplied, defaulting to `mean_se()`
```




```
  ggsave(file="p2.png", width=7, height=16)
```

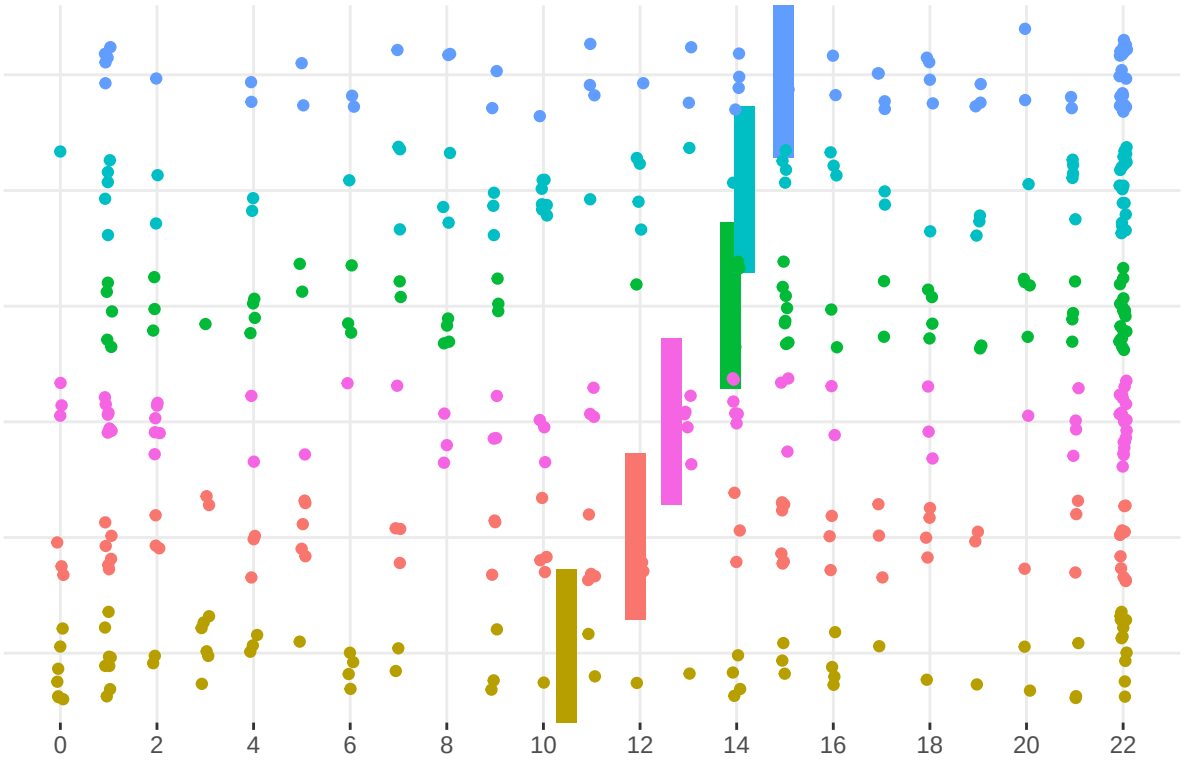
```
## No summary function supplied, defaulting to `mean_se()`
```

Bias Perception

```
lt_RQ2_aggregate %>% ggplot(aes(x = bias.operational, y = reorder(Vis_Type, leastmost), color=Vis_Type)
  stat_summary(geom = "point", shape=108, fun.x = "mean", size=30) +
  geom_jitter(width=0.08) +
  labs(
    title = "",
    x = "",
    y = ""
  ) +
  scale_y_discrete(name ="") +
  theme_bw() +
  theme(legend.key = element_blank(), strip.background = element_rect(colour="white", fill="white")) +
  theme(strip.text.y.left = element_text(angle = 0) ) +
  theme(legend.position = "none") +
  labs(x = "") +
  scale_x_continuous(breaks = seq(0, 30, by = 2)) +
  theme(panel.grid.minor = element_blank(), panel.border = element_blank(), axis.text.y=element_blank()
```

```
## Warning in stat_summary(geom = "point", shape = 108, fun.x = "mean", size =
## 30): Ignoring unknown parameters: `fun.x`
```

```
## No summary function supplied, defaulting to `mean_se()`
```




```
  ggsave(file="p3.png", width=7, height=16)
```

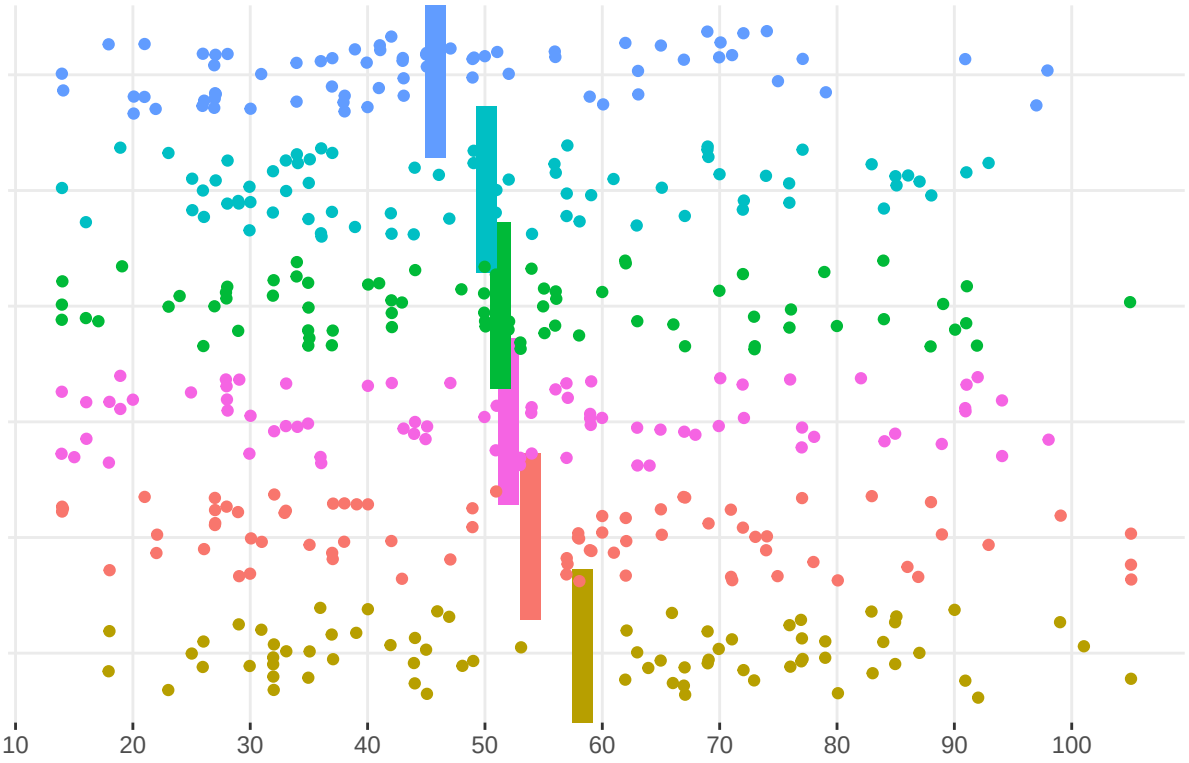
```
## No summary function supplied, defaulting to `mean_se()`
```

Trust

```
lt_RQ2_aggregate %>% ggplot(aes(x = trust, y = reorder(Vis_Type, leastmost), color=Vis_Type)) +
  stat_summary(geom = "point", shape=108, fun.x = "mean", size = 30) +
  geom_jitter(width=0.08) +
  labs(
    title = "",
    x = "",
    y = ""
  ) +
  scale_y_discrete(name ="") +
  theme_bw() +
  theme(legend.key = element_blank(), strip.background = element_rect(colour="white", fill="white")) +
  theme(strip.text.y.left = element_text(angle = 0) ) +
  theme(legend.position = "none") +
  labs(x = "") +
  scale_x_continuous(breaks = seq(0, 100, by = 10)) +
  theme(panel.grid.minor = element_blank(), panel.border = element_blank(), axis.text.y=element_blank()
```

```
## Warning in stat_summary(geom = "point", shape = 108, fun.x = "mean", size =
## 30): Ignoring unknown parameters: `fun.x`
```

```
## No summary function supplied, defaulting to `mean_se()`
```




```
  ggsave(file="p1.png", width=7, height=16)
```

```
## No summary function supplied, defaulting to `mean_se()`
```