8/3/2020

OpenMind Data Report

Prepared by Jack Wychor

Summary

- **Question 1**
  - Intellectual Humility does function on its own as a construct. There is evidence that it consists of a latent two-component structure. These components are tentatively titled “Curiosity” and “Tolerance.”

- **Question 2**
  - Females, Progressives, and Politically polarized individuals (i.e. Very conservative or Very progressive) tended to have higher baseline levels of Affective Polarization. Additionally, Affective Polarization positively correlated with General Belief Superiority (at both pre- and post- measurements), and negatively with Belongingness (post-measurements only).

- **Question 3**
  - Accounting for the random effect of the variation in participants’ scores on Affective Polarization produced a model that fits the data significantly better than a standard linear model. The estimated difference in score attributable to the OpenMind treatment fell between -7.90 and -12.23. The experienced combination of other assessments administered had no effect on this change.

- **Question 4**
  - Age did not change the impact of the OpenMind treatment on Affective Polarization. Males are likely more impacted by the treatment than females. The sample of people gendered as ‘Other’ or `Prefer not to say` was too small to extract generalizable or reliable findings about these groups.

Question 1

The Intellectual Humility (IH) scale consists of 6 questions and is purposed to measure the extent to which participants are willing to change their mind when provided evidence that challenges their core beliefs. Each participant in the OpenMind study (n = 2073) is expected to complete this survey. Participants who did not fully complete the survey were excluded from this analysis (n = 1). Additionally, questions 1, 2, 5, and 6 were reverse coded before analysis. Internal consistency reliability estimates using Cronbach's Alpha as well as item statistics were gathered for the entire scale and recorded in Table 1.

Table 1

Item Statistics for the Intellectual Humility Scale

Question	n	Raw.r	Std.r	r.cor	r.drop	Mean	SD	α
IH1	2072	0.56	0.60	0.46	0.32	5.48	1.27	0.65
IH2	2072	0.59	0.63	0.51	0.37	5.60	1.23	
IH3	2072	0.64	0.54	0.45	0.35	4.69	1.58	
IH4	2072	0.63	0.54	0.46	0.36	5.33	1.49	
IH5	2072	0.57	0.65	0.54	0.40	5.97	0.95	
IH6	2072	0.59	0.67	0.58	0.42	6.01	0.93	

Reliability for the scale falls below the generally accepted 0.70 for the IH scale (α = 0.65). Although not reported, Alpha if item dropped for all questions was lower than the estimated alpha suggesting that poor reliability is not due to poor item performance. In addition, item means show that participants scored relatively high for a 7-point Likert-scale. Standard deviations were slightly low for all questions. Inter-item correlations, however, were acceptable across the board.

Low internal-consistency reliability may be indicative of multiple latent factors within the responses. A principal component analysis (PCA) was used to test this. First, Eigenvalues for models with components from 1 to 6 were plotted on a scree plot. Figure 1 shows a 2-component model fit the data best. Based on the PCA results in Table 2, there is a very clean split between component 1 (questions 1, 2, 5, and 6) and component 2 (questions 3 and 4). Note that this may be at least partially due to question wording, as this is the same split for questions that were reverse coded. Given the verbiage for questions in component 1, it was given the title "Curiosity" while the wording for questions in component 2 suggest "Tolerance."

Figure 1. Scree Plot for IH scale responses

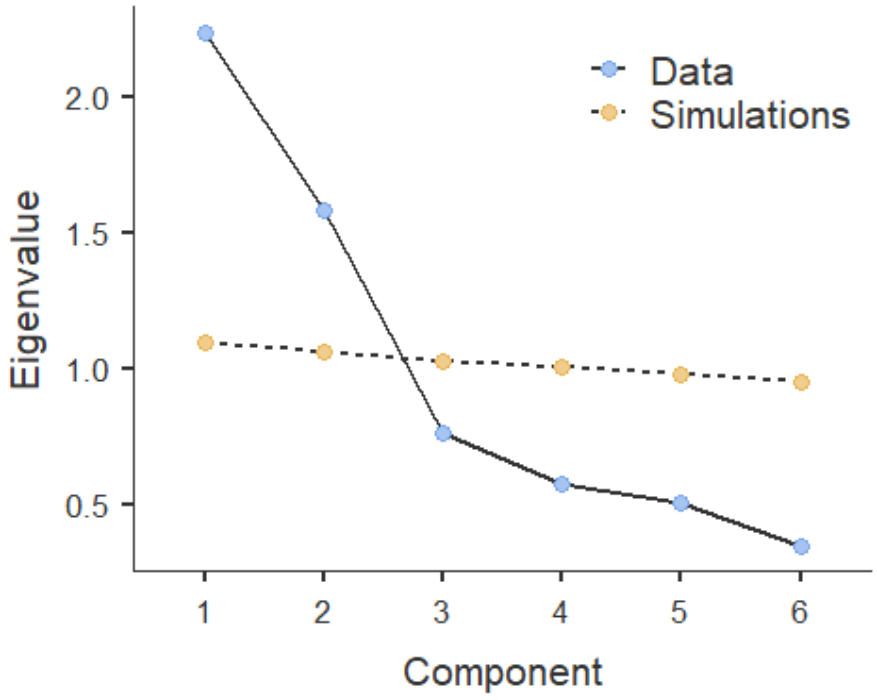


Table 2

2-Component Principal Components Analysis of IH Scale Responses

Question	Component 1 (Curiosity) Loading	Component 2 (Tolerance) Loading	Uniqueness
IH1	0.72		0.48
IH2	0.73		0.47
IH3		0.91	0.18
IH4		0.91	0.18
IH5	0.72		0.46
IH6	0.76		0.41

Note. Loadings < 0.65 were left blank. Cumulative variance explained = 63.53%.

Cronbach's Alpha narrowly reached acceptable levels for component 1 (α = 0.70) and moderately strong for component 2 (α = 0.79) shown in Tables 3 – 4. **This is evidence that a two-component structure for the construct of Intellectual Humility consisting of Curiosity and Tolerance as sub-components**. If researchers plan to continue with this structure, it is recommended that more questions be added as both scales have a relatively low number of questions. This will also increase the reliability of both scales.

Table 3

Item Statistics for Component 1 (Curiosity)

Question	n	Raw.r	Std.r	r.cor	r.drop	Mean	SD	α
IH1	2072	0.76	0.72	0.56	0.48	5.48	1.27	0.70
IH2	2072	0.76	0.73	0.58	0.50	5.60	1.23	
IH5	2072	0.69	0.73	0.60	0.48	5.97	0.95	
IH6	2072	0.72	0.76	0.65	0.53	6.01	0.93	

Table 4

Item Statistics for Component 2 (Tolerance)

Question	n	Raw.r	Std.r	r.cor	r.drop	Mean	SD	α
IH3	2072	0.91	0.91	0.73	0.65	4.69	1.58	0.79
IH4	2072	0.9	0.91	0.73	0.65	5.33	1.49	

Question 2

When analyzing the relationship between demographic variables and the measured construct Affective Polarization (AP), it is important to consider the types of variables being compared. A correlation matrix was made to demonstrate the relationship between pre- and post- measures of AP and all other continuous variables measured. It is important to note that only pairwise complete observations were used as not every participant was assessed on all variables. On top of this, Age and ID were the only continuous demographic variables, the remaining variables were captured in assessments. Correlations were considered noteworthy if they exceeded $|r| > 0.20$. Table 5 shows there is a strong relationship between AP and General Believe Superiority (GBSS) scores (Pre- and Post-) and belongingness (Post- only).

A factorial ANOVA was used to compare the effects of categorical demographic variables on AP pre- scores (Note that Age is the only demographic variable in the correlation matrix). AP post- scores were ignored in this analysis as measuring these would effectively represent the interaction between the variable *and* the OpenMind treatment. Each Race/Ethnicity was treated as a separate dummy-coded variable because they were not mutually exclusive for each participant. The resulting model in Table 6 shows that Gender ($F(3,494) = 7.05$, $p < 0.01$) and Political Affiliation ($F(5,494) = 30.56$, $p < 0.01$) showed significant differences between levels. Further investigation (through planned contrasts) found Females were significantly higher than Males on AP ($t(515) = 2.87$, $p < 0.01$), Very politically left/right individuals were significantly higher than less polarized individuals ($t(515) = 8.03$, $p < 0.01$), and Progressives are

higher on AP than Conservatives ($t(515) = 4.42$, $p < 0.01$). Figure 2 provides a graph comparing means for each level of Gender and Political Affiliation. **In sum, the following demographics had significantly higher AP scores:**

1. **Females**
2. **Progressives**
3. **Politically polarized individuals**

Additionally, AP scores correlated highly with GBSS scores (Pre- and Post-), and Belongingness scores (Post- only).

Table 5

Correlations between AP Scores (pre- and post-) and Continuous Variables

Variable	AP Pre-	AP Post-
ID	0.11	0.15
Age	-0.01	0.02
IH1Pre	0.00	-0.07
IH2Pre	0.13	0.14
IH3Pre	-0.11	-0.05
IH4Pre	-0.08	-0.07
IH5Pre	-0.09	-0.01
IH6Pre	-0.01	-0.05
GMPre	0.06	-0.05
GMPost	-0.10	-0.18
SEPre	0.07	0.05
SEPost	0.03	0.01
AffPol2Pre		0.76*
AffPol2Post	0.76*	
AnxietyPre	0.10	0.13
AnxietyPost	0.06	0.09
GBSSPre	0.33*	0.21*
GBSSPost	0.32*	0.46*
BelongPre	-0.09	-0.18
BelongPost	-0.12	-0.25*

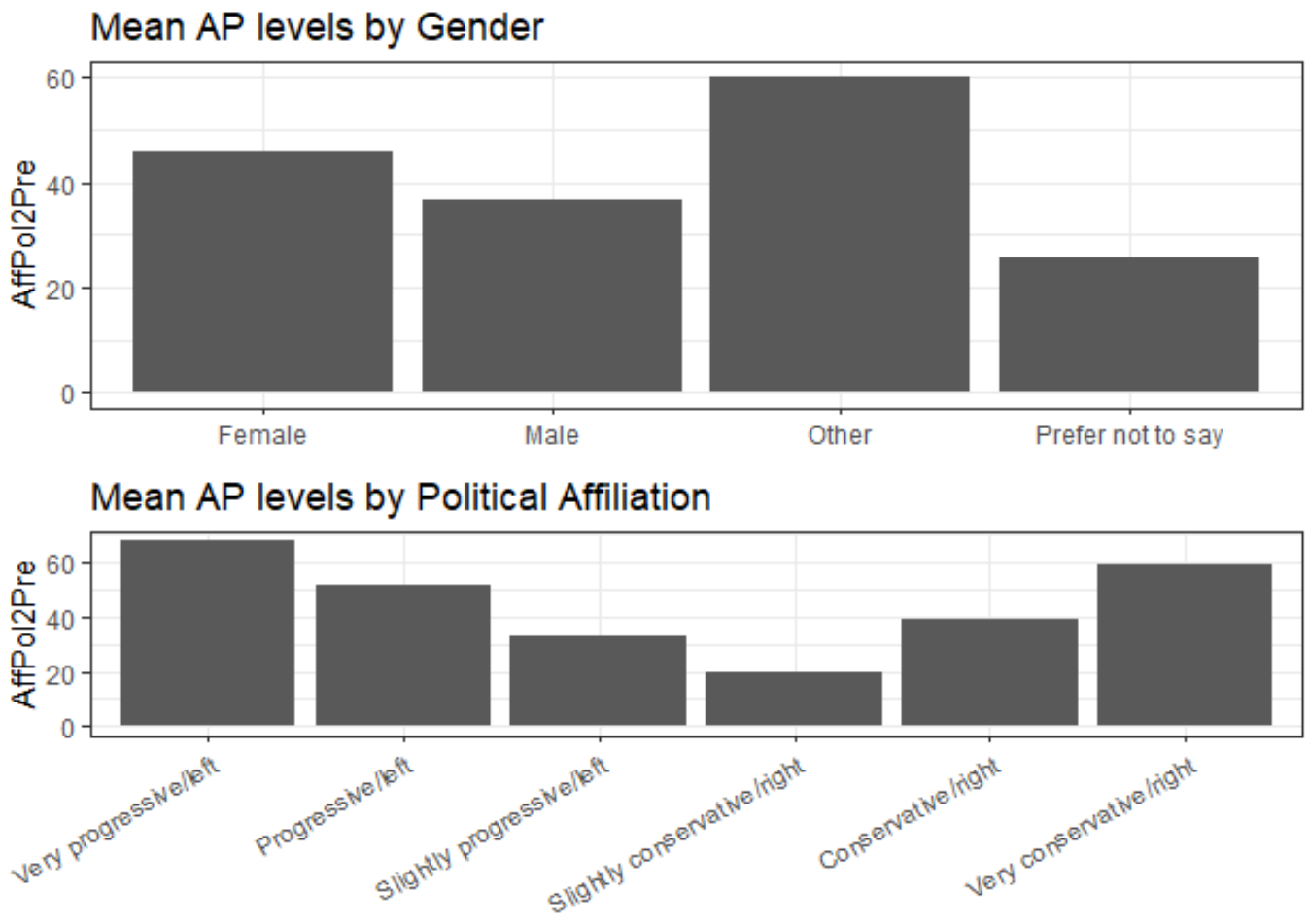
Note. * = $|r| > 0.20$.

Table 6

ANOVA Results between AP Pre- Scores and Categorical Variables

Variable	DF	Sum Sq	Mean Sq	F Value	P Value
Gender	3	12302.56	4100.85	7.05	< 0.01*
Political Orientation	5	88917.4	17783.48	30.56	< 0.01*
Religious Affiliation	13	6816.49	524.35	0.9	0.55
Hispanic or Latino	1	18.79	18.79	0.03	0.86
African American or Black	1	256.82	256.82	0.44	0.51
South Asian	1	605.21	605.21	1.04	0.31
East or Southeast Asian	1	1757.59	1757.59	3.02	0.08
White or Caucasian	1	363.71	363.71	0.62	0.43
Other	1	746.94	746.94	1.28	0.26
American Indian or Alaska Native	1	161.12	161.12	0.28	0.60
Prefer not to Say	1	26.62	26.62	0.05	0.83
Residuals	494	287514.05	582.01		

Figure 2. Mean AP Scores by Demographic Group



Question 3

A linear mixed effects model was used to analyze the fixed effect of the OpenMind treatment on AP Post- assessment scores compared to Pre- scores while accounting for the random effects of participant score variance and the combination of treatments given. Treating participant score as a random effect is a common method for handling random variance in performance for within-subjects designs. Because the combination of assessments that each participant could take was randomly distributed, it was also treated as a random effect. This was done by recording the combination of assessments that participant took in a 6-digit binary string (i.e. 000111) where each digit represented whether the assessment in that index was taken (1) or not (0). For reference, the order of the assessments and their index was as follows: Growth Mindset (1), Self-Esteem (2), AP (3), General Anxiety (4), GBSS (5), and Belonging (6). Therefore, only codes that had a "1" in the third digit of the code were included (i.e. 001000, but not 110010) in this linear model because AP is the dependent variable and its index is 3.

To analyze whether the treatment code was a significant random variable, a mixed linear model (Model 1) including variance due to participant and variance due to treatment as random variables was tested against a model (Model 2) including only the former as a random variable i.e.:

```
Model1: AffPol2 ~ time + (1 | ID) + (1 | treatcode)
Model2: AffPol2 ~ time + (1 | ID)
```

The models did not significantly differ ($\chi^2 = 0$, $p = 1$). Model 2 should be used over the Model 1 because it has less terms. Next, Model 2 was then tested against a standard linear model with no random effects i.e.:

```
Model2: AffPol2 ~ time + (1 | ID)
Model3: AffPol2 ~ time
```

Model 2 fit the data significantly better than Model 3 ($\chi^2 = 247.71$, $p < 0.01$). We can therefore infer that including the random effect in the model is significantly better than using only the fixed effect. **The estimated difference attributable to the OpenMind treatment on AP was between -12.23 and -7.90 with 95% confidence. Because AP was measured from 0 - 100, this is an average of a 10% reduction.**

Question 4

A linear mixed effects model including the effect Age and Gender of participants on the difference in Pre- and Post- AP scores was fit using the following formula:

```
Model4: AffPol2 ~ time + time:age + time:gender + (1 | ID)
```

Model 4 tests the main effects of Time plus the interaction effects of Time:Age and Time:Gender. Main effects of Age and Gender were avoided as we are only interested in the effect these variables have on the *effect* of OpenMind on AP scores (Time). In addition, Time is the first variable because R uses Type 1 sums of squares to calculate the variance explained for predictors (i.e. each successive predictor cannot explain the variance that is explained by a prior predictor). Finally, the three-way interaction between Time, Age, and Gender was avoided for sake of having a more explainable model. Figure 3 provides a summary of the output results for Model 4. The effect of OpenMind on AP was equal for all ages. In relationship to Females (n = 319), Males (n = 198) likely scored slightly lower with a low spread at 95% confidence (-5.72 to 0.13), `Prefer not to say` (n = 5) decreased even more than Males-- but with a larger spread at 95% confidence (-23.71 to 7.40), and Other (n = 2) had the largest spread of expected change at 95% confidence (-10.83 to 34.76). **It is very likely Males are more affected by the OpenMind treatment than Females**. The findings for Other and `Prefer not to say` have poor generalizability and reliability due to small sample size and this is reflected in their massive confidence intervals.

Figure 3. Effects of OpenMind and Interactions with OpenMind with Gender and Age on AP