Data Quality Summary

For Emotional States Prediction Projects

Sicong Zhao @ Duke University
sicong.zhao@duke.edu

Data Quality Summary

I went through the box plots of 208 features and detected following columns are suspicious

(1) Affect Valuation Index (AVI) (16 features)

The problem could be, in ROO experiment, we are using average value, but in DND experiment, we are using sum.

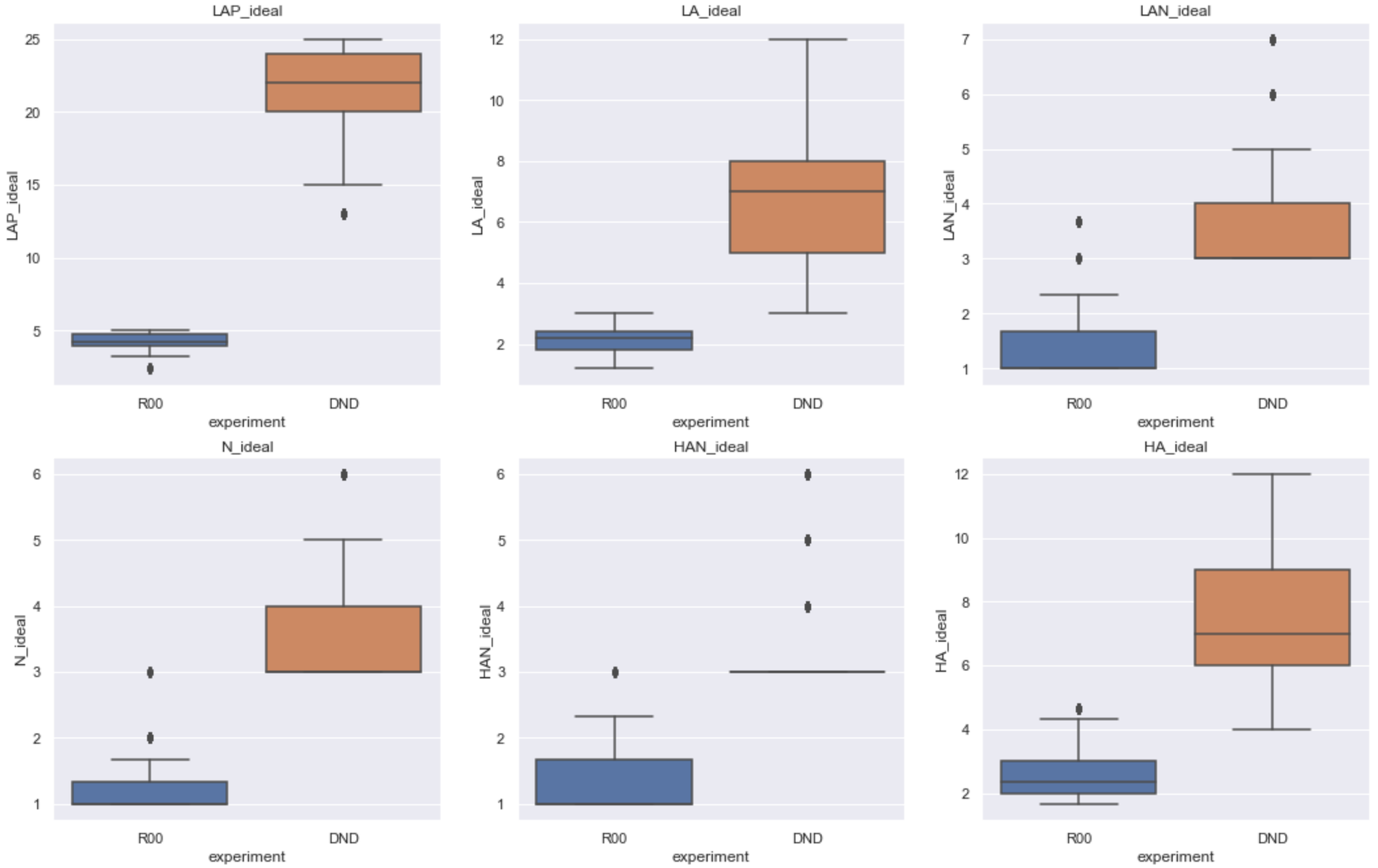
(2) Personality Measurement (5 features)

1.Neuroticism: Scale not aligned
2. Extraversion: Scale not aligned
3. Conscientiousness: Scale not aligned
4. Agreeableness: Completely missing in DND
5. Openness: Completely missing in DND

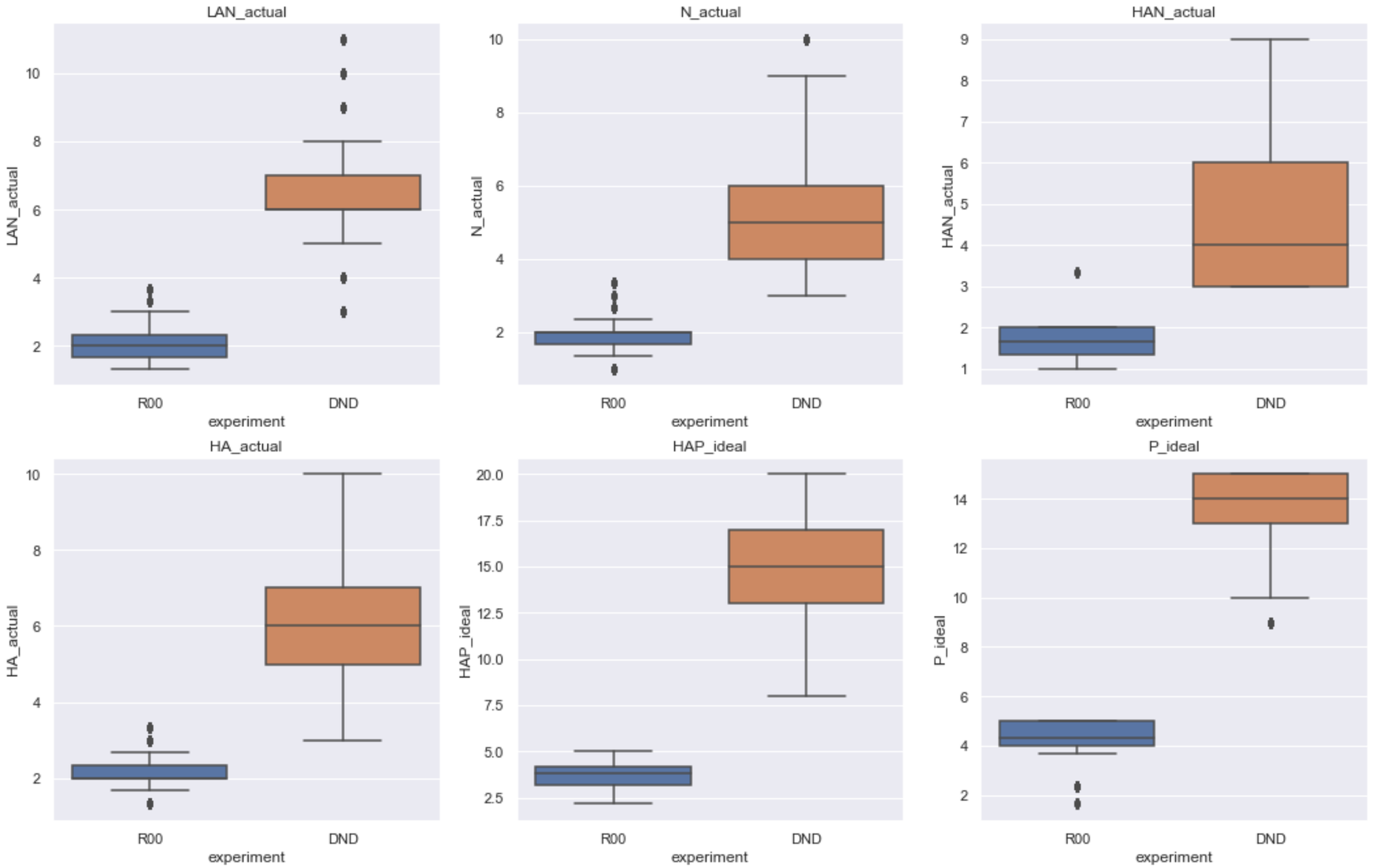
(3) Others (6 features)

1. SBQ
2. FTP
3. NS_total
4. BIS.5
5. BAS_D
6. BAS_RR,

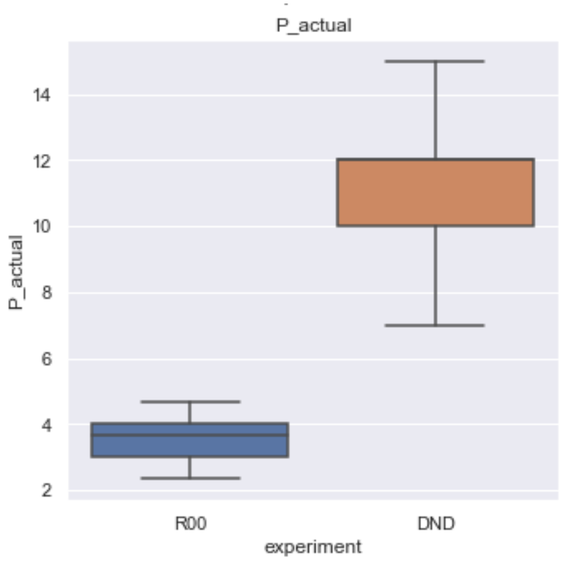
1.AVI – Part1

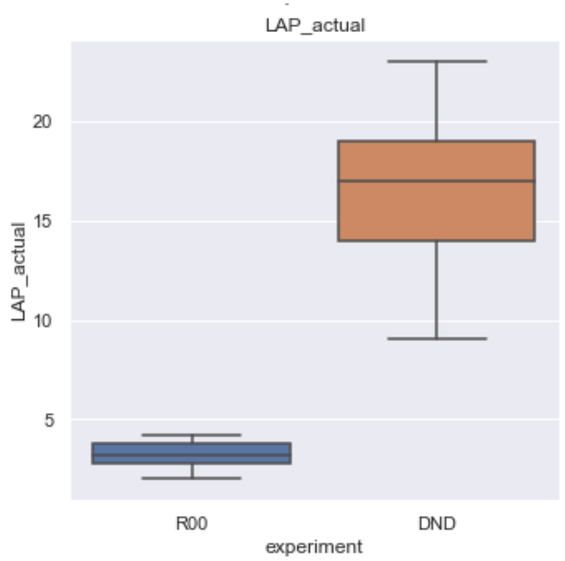

LAP_ideal
LAP_ideal
25
20
15
10
5
R00
DND
experiment
LA_ideal
LA_ideal
12
10
8
6
4
2
R00
DND
experiment
LAN_ideal
LAN_ideal
7
6
5
4
3
2
1
R00
DND
experiment
N_ideal
N_ideal
6
5
4
3
2
1
R00
DND
experiment
HAN_ideal
HAN_ideal
6
5
4
3
2
1
R00
DND
experiment
HA_ideal
HA_ideal
12
10
8
6
4
2
R00
DND
experiment

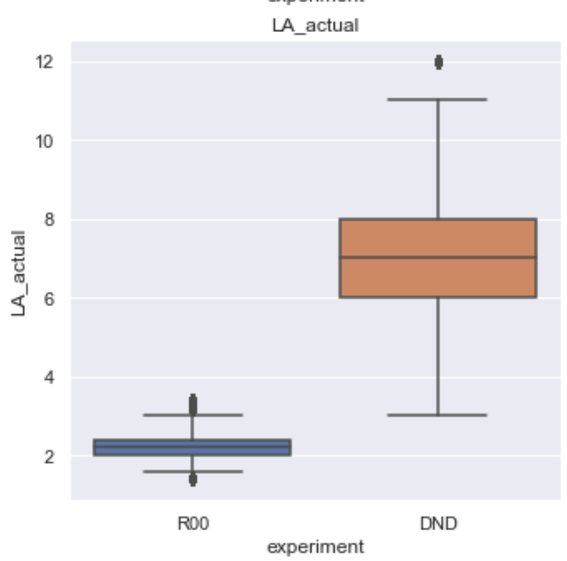
1.AVI – Part2

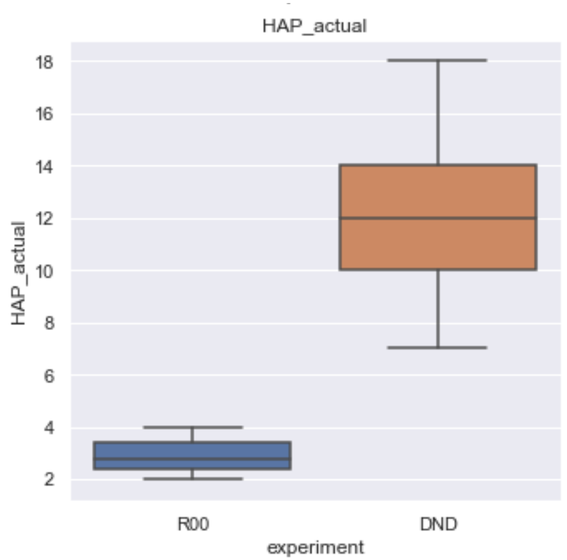

LAN_actual
LAN_actual
experiment
R00
DND
N_actual
N_actual
experiment
R00
DND
HAN_actual
HAN_actual
experiment
R00
DND
HA_actual
HA_actual
experiment
R00
DND
HAP_ideal
HAP_ideal
experiment
R00
DND
P_ideal
P_ideal
experiment
R00
DND

1.AVI – Part3

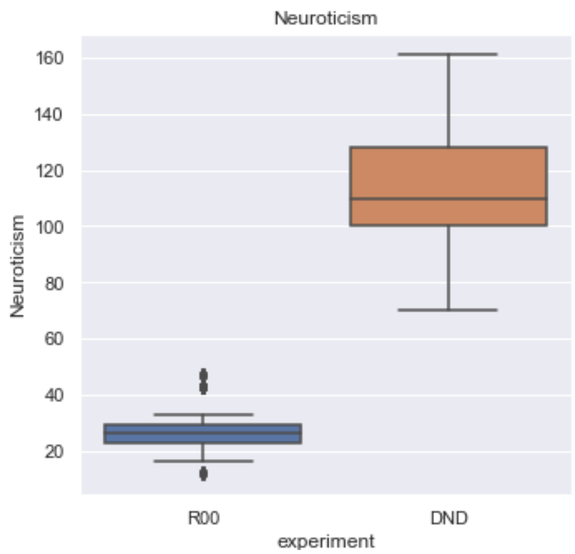

P_actual
P_actual
14
12
10
8
6
4
2
R00
DND
experiment

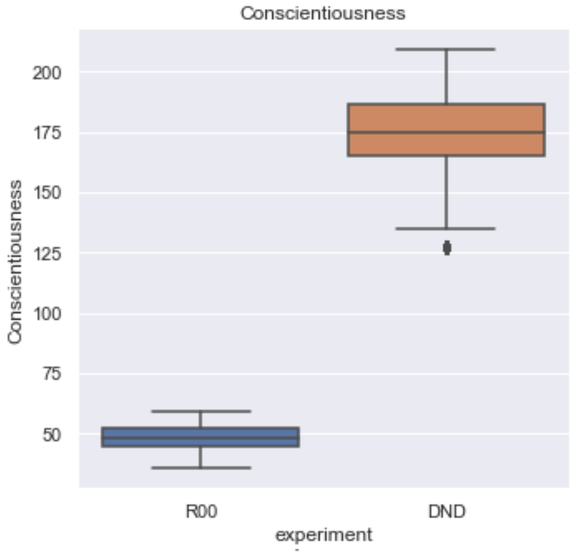

LAP_actual
LAP_actual
20
15
10
5
R00
DND
experiment

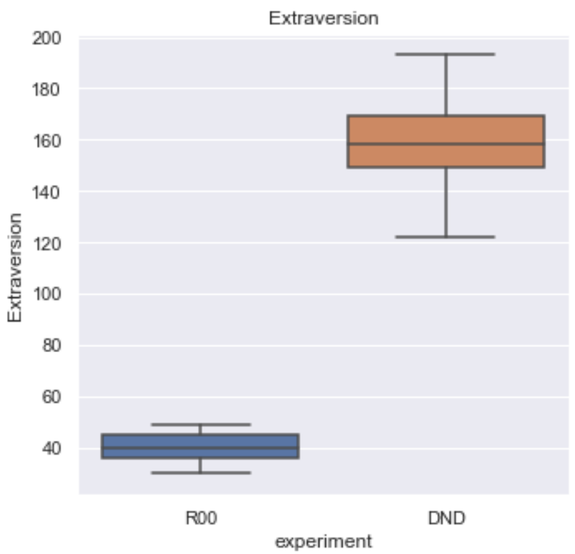

LA_actual
LA_actual
12
10
8
6
4
2
R00
DND
experiment

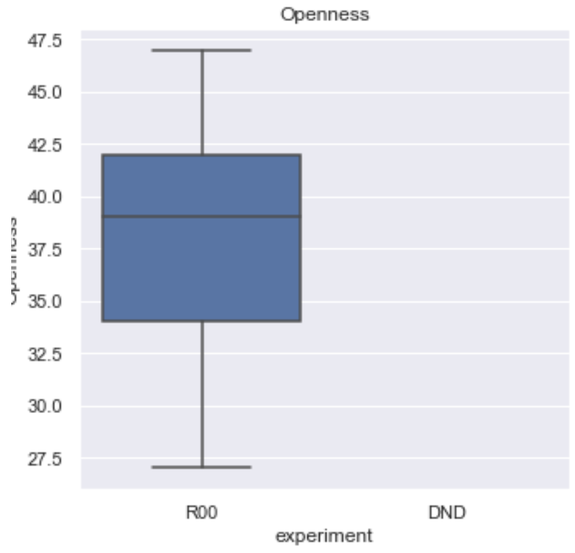

HAP_actual
HAP_actual
18
16
14
12
10
8
6
4
2
R00
DND
experiment

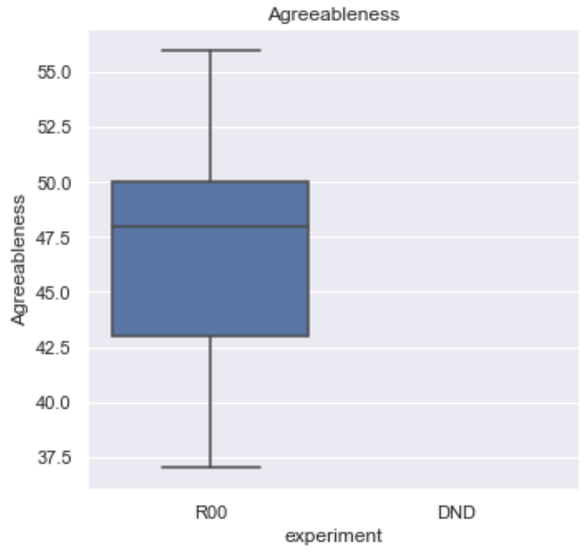
2. Personality



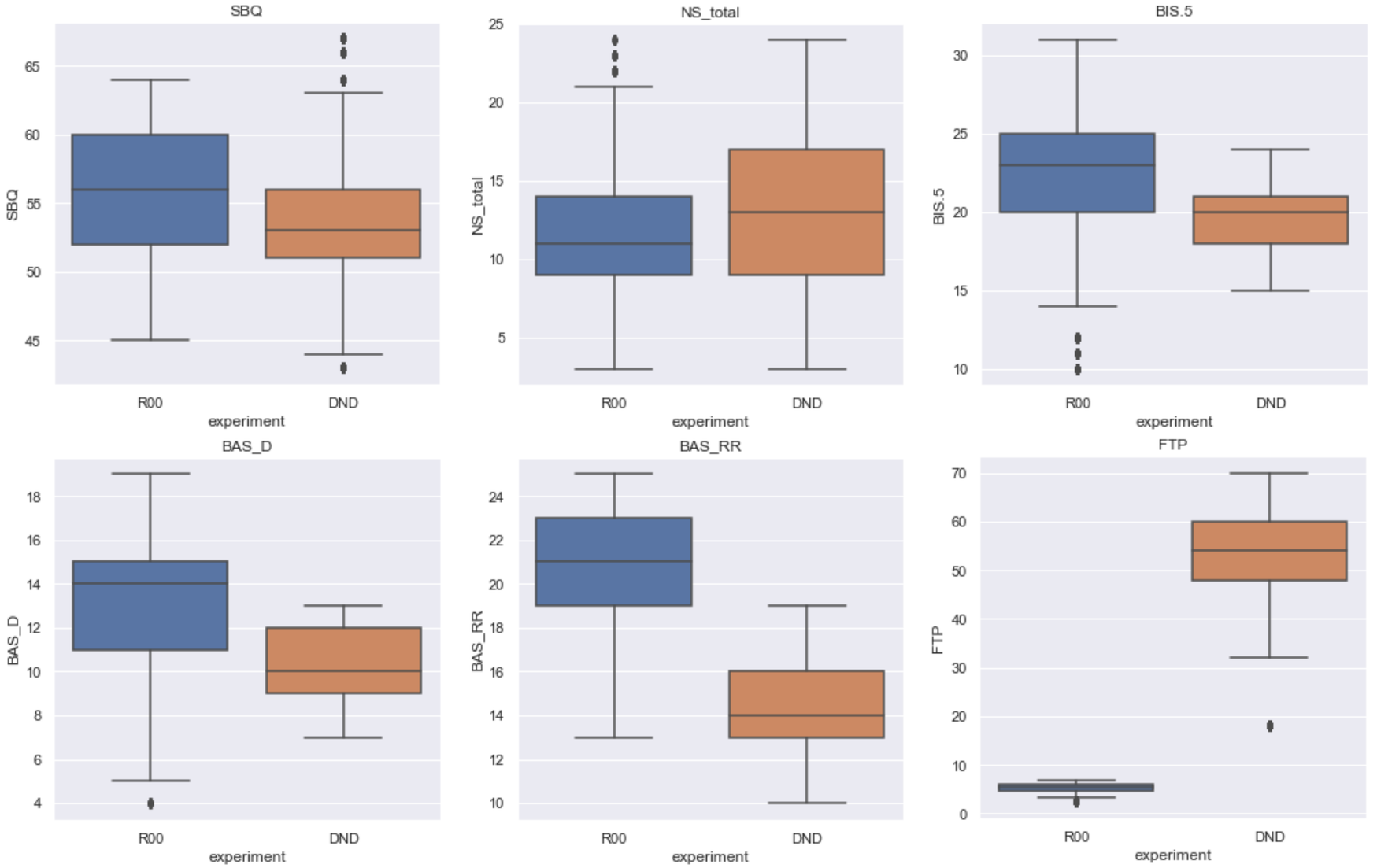








3. Others


SBQ
SBQ
65
60
55
50
45
R00
DND
experiment
NS_total
NS_total
25
20
15
10
5
R00
DND
experiment
BIS.5
BIS.5
30
25
20
15
10
R00
DND
experiment
BAS_D
BAS_D
18
16
14
12
10
8
6
4
R00
DND
experiment
BAS_RR
BAS_RR
24
22
20
18
16
14
12
10
R00
DND
experiment
FTP
FTP
70
60
50
40
30
20
10
0
R00
DND
experiment

4. Solutions

(1) Affect Valuation Index (AVI) (16 features)

(2) Personality Measurement (5 features) Z-Score within study

1.Neuroticism: Scale not aligned
2. Extraversion: Scale not aligned
3. Conscientiousness: Scale not aligned
4. Agreeableness: Completely missing in DND
5. Openness: Completely missing in DND

(3) Others (6 features)

1. SBQ
2. FTP (sun / average - z-score)
3. NS_total
4. BIS.5
5. BAS_D
6. BAS_RR,