

Perceived hatefulness

Hate speech

Offensive

Neither

Mean score (Max value = 1)

No consequence

0.15

0.67

0.96

Remove post

0.76

0.26

0.02

Ban user

0.57

0.11

0.01

Legal action

0.51

0.11

0.02

Job

0.40

0.09

0.02