对“Are You for Real? Perceptions of Authenticity Are Systematically Biased and Not Accurate”的可重复性研究

摘 要 人们能够准确地感知到他人是真实的吗？本研究的目标是在随机分配的被试中比较自我评价和他人评价的真实性。假设为：自我评价和他人评价的真实性之间的关系不显著，结果也验证了假设。另外，该研究还测试了感知真实性的偏差。首先比较了他人评价的真实性和自我评价的真实性，结果发现他人评价的真实性要高于自我评价的真实性。另外还发现了评分者自身的真实性可以预测他们对他人真实性的评级。总体来说，我们发现人们不能准确的辨别他人是否是真实的。

关键词 自我评价，他人评价，真实性

1. 引言

1.1研究的简述

在研究二中，研究者直接测试感知真实性的准确性时，发现在随机分配的两组成年学生中，自我和他人评价的真实性之间没有显著的相关性。此外，还发现感知真实性在两个方面存在偏差:(1)他人评价的真实性与自我评价相比表现出积极的偏差；(2)他人评价的真实性受到评价者自身真实性的偏差。

1.2研究的假设

假设1：自我评价和他人评价之间的关系是不显著的；假设2：他人评价的真实性高于自我评价的真实性；假设3：评价者自己的真实性会预测他们对其他人的真实性评分。

1.3研究的方法

当学生进入MBA课程时，我们进行了调查1，统计人口学和自我评价状态真实性这两个变量。第二周开始进行调查2，被试被要求完成他们自己的特征自我评估(自评特质真实性和个性)。在四周后，也就是第六周，被试被要求在调查3中评估小组成员的特征(他人评价的特质真实性和个性)。被试有2周的时间来完成对小组成员的评估(即调查4前后各1周)。6周结束时进行最后一次调查，即调查4，被试被要求完成状态真实性、行为真实性和被了解的感觉这三项自评任务。在这个小组内，被试需要学习第一学期的所有课程，并完成一些评分任务，小组成员花了大量的时间在一起。并被要求开会制定规范、规则和责任章程。因此，参与者基本上是在同一时间相互参与的。其次，互动主要是在相同或相似的环境中进行的(教室或学校设施)。

研究二主要对自我评定的特质真实性和他人评定的特质真实性进行了比较。然而，被试在估计目标的特质真实性时，有可能依赖于真实性的替代表达(例如，真实性感觉)。为了控制这一点，研究还要求被试报告他们的真实感受，例如：他们是否觉得自己能和队友真诚地相处，以及他们是否觉得队友知道他们的真实身份。然后，对这些额外形式的自我真实性与他人评价的特质真实性的一致性进行测试。

我们使用一种独特的人际环境测试了真实性评级的准确性和偏差:随机分配的学生团队彼此之间几乎没有互动。对这些群体的随机分配减轻了在个人和职业关系中通常出现的选择问题。我们对这些人进行了为期6周的跟踪调查，在多个时间点对他们进行了调查。参与者对自己的真实性和小组中其他人的真实性进行评估。这使我们能够直接测试自我和他人对真实性的评价之间的一致性，以及测试真实性判断中的偏差。真实性是一个多方面的结构，本研究从特质角度和状态角度来考虑。

在自我评分这一维度下，本研究探索了：“特质真实性”、“状态真实性”、“被了解的感觉这一变量”以及“人格”这四个变量。对“特质真实性”这一变量，本研究采用从真实人格量表(Wood et al, 2008)修改的项目和新创建的项目组合来评估他们的真实性。对“状态真实性”这一变量，研究通过项目开始时的调查1和6周后的调查4，对被试真实性的感觉进行了评级，并对部分项目进行反向编码。对“被了解的感觉” 这一变量，通过“我的团队成员知道我真正是谁”和“我在大学的朋友知道我真正是谁”这两个问题来对被试进行测量。对于“人格”这一变量，研究使用十项人格量表(Ehrhart et al, 2009)对被试的五大人格进行评分。之所以使用这一变量，第一是因为他人对人格特质的评价是相对准确的(Ambady & Rosenthal, 1992)。因此，可以通过比较自我评价和他人评价的个性，为真实性的准确性创建基线。第二，可以使用人格特征作为控制变量来检验模型的稳健性。第三，可以通过人格特质来探索被试的真实性评级。

在他人评分这一维度下，本研究探索了：“特质真实性”、“评定等级的真实性”、“真实性错误”以及“人格”这四个变量。每个提交了自我评价的被试都被要求完成对团队中每个成员的评价。对“特质真实性”这一变量，每个被试被其团队成员根据他们对基于特质的自我评价中相同的四个项目的真实性的感知来打分。对“评定等级的真实性”这一变量，从被试的自评调查中匹配被试的真实性。对“真实性-错误”这一变量，本研究将该变量的得分作为自我和他人评价真实性之间差异的指标。该变量是通过从其他评定的特质真实性中减去自评特质真实性来估计的，正的真实性-错误分数表明他者评价高于自我评价的特质真实性。真实性-错误分数为零，表明自我评级和他人评级基本不存在差异。负的真实性-错误分数表明他人评定的特质真实性低于自我评定的特质真实性。对于“人格”这一变量，研究使用十项人格量表(Ehrhart et al, 2009)要求被试对目标任务的五大人格进行评分。最后，对被试对目标人物的熟悉程度进行测量。

1.4研究的结果

该研究使用普通最小二乘(OLS)回归模型，在“真实性”“感觉真实性”“行为真实性”“被了解的感觉”四个项目上，比较了自我评定的特质真实性和他人报告的特质真实性的差异。首先，我们发现自我评价的特质真实性不能预测他人评价的特质真实性*(β* = 0.04, *SE* = 0.03, 95% *CI* = [- 0.01, 0.09]，*p* = .132)。其次，在第一次和第四次调查中，由他人评定的特质真实性与自我评定的状态真实性也无关(调查1: *β* = 0.01, *SE* = 0.03, 95% *CI* = [- 0.04, 0.06]，*p* = .608;调查4: *β* = 0.02, *SE* = 0.03, 95% *CI* = [-0.03, 0.08]，*p* = .348)。接着，由他人评价的特质真实性也与被试对自己能够与队友真诚合作的感知无关(*β* = 0.00, *SE* = 0.03, 95% *CI* = [- 0.05, 0.05]，*p* = .980)。最后，由他人评价的特质真实性也无法预测被试对队友认识真实的他们这一维度上的评分(*β* = - 0.04, *SE* = 0.03, 95% *CI* = [- 0.09, 0.01]，*p* = .142)。综上所述，四种结果共同表明他人评价的真实性与自我评价的真实性不相关。

为了稳健性，我们使用非信息先验估计了自我评价和他人评价真实性之间关系的贝叶斯回归，以获得回归系数的可信区间(Wagenmakers et al，2016)。研究发现，他人评定特质真实性系数对自我评价特质真实性系数的95%可信区间为[-0.05,0.07]，包含0。此外，我们使用实际等效区域(ROPE)方法测试了两者的无相关关系(Kruschke & Liddell, 2018)。研究发现，在自我评价和他人评价的特质真实性之间的关系上，HDI和ROPE完全重叠(100.00%)。这表明零假设没有被拒绝，即自我评价真实性与他人评价的真实性之间没有显著的相关关系。

为了检验关系的哪个方面解释了真实性评级的差异，我们进行了SRM(Social-relations modeling) (Kenny & La Voie, 1984)分析，它将自我和他人评级的方差分解为：做出判断的人、导致判断的目标特征以及自我-他人配对特有的这三个部分。此外，SRM在控制群体成员关系的条件下，还提供了他人评定与自我评定的部分相关性(即，真实性与被认为真实之间的相关性是什么?)，以及被试评分与他们给出的他人评价的相关性。

SRM的结果表明，目标间的差异，解释了约6%的方差(*SE* = 0.02, *p* = 0.003)。另一方面，被试间差异解释了真实性评分中41%的方差(*SE* = 0.06, *p* < .001)，自我-他人关系的差异解释了真实性评分中45%的方差(*SE* = 0.03, *p* < .001)。这表明，真实性评级的大部分差异与做出评分的被试之间的差异以及自我-他人不同关系之间的差异有关。

在控制群体成员的情况下，自评特质真实性与他人评价特质真实性的偏相关不显著，*r* = 0.12, *t*(167) = 1.53, *p* = 0.129。然而，在控制了群体成员关后，评价者的真实性与他们对他人真实性的感知之间存在显著的偏相关，*r* = 0.34, *t*(167) = 4.69, *p* < 0.001。

为了这个样本中的评分者并不能准确地判断他们的目标这一可能性，研究对自评人格特质和他评人格特质再次进行了一系列多层次模型，两者的正相关关系表明，自评人格特质得分对他评人格特质得分具有正向预测作用，这一结果表明评价者的人格评价是准确的。

研究发现，他人评定的特质真实性(*M* = 5.84)显著高于自我评定的特质真实性(*M* = 5.27;*SD*= 0.57,95% *CI* = [0.43, 0.71])，*t*(315) = 7.98, *p* < 0.001, *Cohen’s d* = 0.79。也就是说，参与者更有可能认为其他人比他们自己更真实。

同时，研究还发现被试的特质真实性与其对其他个体的真实性评分之间存在显著的正相关关系(*β* = 0.23, *SE* = 0.03, 95% *CI* = [0.17, 0.29]，p < .001)。当控制了评分者的个性、人口统计学特征和对目标的熟悉程度时，这种效应仍然存在(*β* = 0.20, *SE* = 0.03, 95% *CI* = [0.13, 0.26]，*p* < .001)。

为了解释评分者的偏见是由于项目的相似性，而不是他们自己的真实性这一观点，研究控制了其他真实性变量(状态、行为真实、被了解的感觉)与真实性评级之间的关系。我们发现调查1和调查4这两个时间点的状态真实性对其他个体的真实性判断有显著的正向影响(调查1: *β*= 0.12, *SE* = 0.03, 95% *CI* = [0.06, 0.18]，*p* < .001;调查4: *β*= 0.11, *SE* = 0.03, 95% *CI* = [0.04, 0.17]，*p* < .001)。

研究还发现被试认为自己的行为是真实的可以正向预测被试对他人的真实性判断(*β*= 0.10, *SE* = 0.03, 95% *CI* = [0.04, 0.17]，*p* = 0.002)，同样也可以正向预测被试对群体成员被了解的感受的真实性判断(*β* = 0.19, *SE* = 0.03, 95% *CI* = [0.12, 0.25]，*p* < 0.001) 。然后，研究以被试的真实性作为符号差异真实性-错误得分的预测因子。被试的特质真实性预测了正的真实性-错误评分(*β* = 0.24, *SE* = 0.03, 95% *CI* = [0.18, 0.30]，*p* < .001)。也就是说，真实的被试认为他评真实性相较于自评真实性更真实。

最后，研究为了探索那些被试更有可能被评为真实的，对目标的人口统计学和人格变量进行了探索性分析，并在评分水平上聚类，。结果没有发现存在任何个体差异可以预测他人特质真实性的评价 (*ps* > .11)。

1. 研究复现思路及流程

基础的数据清理和预处理

加载需要的包，并读取原始数据文件：

# Load required packages

library(dplyr)

library(report)

library(psych)

library(patchwork)

library(ggplot2)

library(apaTables)

library(lme4)

library(lmerTest)

library(rstanarm)

library(TripleR)

library(bayestestR)

# Reading in Data

mydata = read.csv("AYFR\_Study2\_import.csv")

变量注释：'.a'：他人评价自己；'.o'：自己评定他人；'obs\_'：自己评价自己

状态真实性：x.1 = Survey 1; x.2 = Survey 4。

# Note: Participants in the data set are both actors (one's being rated; typically noted with '.a' at the end) and observers (one's doing the rating; typically noted with '.o' at the end)

# In the code, "actor" variables or variables without a prefix indicate self-rated or variables answered in regards to the self; "observer" variables indicate other-rated or variables answered in regards to someone else.

# For biases, I use observer's self-rated variables as predictors of their ratings of others.

# These variables begin with 'obs\_' to indicate that they are self-rated by the observers.

# (x.1 = Survey 1; x.2 = Survey 4)

对研究二中的调查1和调查4中的“状态真实性”中需要反向积分的问题进行反向计分，并求均值

# State authenticity

#survey1

mydata$cn1.r = 8 - mydata$cn1

mydata$cn2.r = 8 - mydata$cn2

#survey4

mydata$cn1.2.r = 8 - mydata$cn1.2 #我觉得我假装自己是别人

mydata$cn2.2.r = 8 - mydata$cn2.2 #我自己虚伪

mydata$fas.1 = (mydata$cn1.r + mydata$cn2.r)/2 #survey1两个问题的平均值

mydata$fas.2 = (mydata$cn1.2.r + mydata$cn2.2.r)/2 #survey4两个问题的平均值

#survey1

mydata$obs\_cn1.r = 8 - mydata$obs\_cn1

mydata$obs\_cn2.r = 8 - mydata$obs\_cn2

#survey4

mydata$obs\_cn1.2.r = 8 - mydata$obs\_cn1.2

mydata$obs\_cn2.2.r = 8 - mydata$obs\_cn2.2

mydata$obs\_fas.1 = (mydata$obs\_cn1.r + mydata$obs\_cn2.r)/2 #survey1两个问题的平均值

mydata$obs\_fas.2 = (mydata$obs\_cn1.2.r + mydata$obs\_cn2.2.r)/2 #survey4两个问题的平均值

对研究二中的调查4的“行为真实性”需要反向积分的问题进行反向计分，并求均值

# Acting authentically

#Q1：有时我觉得我和同学们在一起时不能做我自己

#Q2：我可以和我的队友们坦诚相待

#Q3：我是那种喜欢和同学们在一起的人

mydata$acting.authentic\_1.r = 8 - mydata$acting.authentic\_1

mydata$acting.auth = (mydata$acting.authentic\_1.r + mydata$acting.authentic\_2 + mydata$acting.authentic\_3)/3

mydata$obs\_acting.authentic\_1.r = 8 - mydata$obs\_acting.authentic\_1

mydata$obs\_acting.auth = (mydata$obs\_acting.authentic\_1.r + mydata$obs\_acting.authentic\_2 + mydata$obs\_acting.authentic\_3)/3

对研究二中调查4中的“被了解的感觉”变量的问题求均值

# Feeling known 被了解的感觉

#Q1：我的团队成员知道我真正是谁

#Q2：我的大学朋友知道我真正是谁

mydata$known = (mydata$known\_1 + mydata$known\_2)/2

mydata$obs\_known = (mydata$obs\_known\_1 + mydata$obs\_known\_2)/2

对研究二中大五人格得分、特质真实性进行反向计分和求平均值

# Personality

#extra\_1.a：Extraverted外向, enthusiastic热情

#extra\_2.a：Reserved矜持, quiet安静

#agree\_1.a：critical quarrelsome激烈地争吵

#agree\_2.a：sympathetic同情, warm温暖

#conc\_1.a：Dependable可靠, self-disciplined自律

#conc\_2.a：Disorganized无秩序的, careless粗心大意

#neuro\_1.a：Anxious焦虑, easily upset容易沮丧

#neuro\_2.a：Calm冷静, emotionally stable情绪稳定

#open\_1.a：Open to new experiences接受新体验, complex复杂

#open\_2.a：Conventional传统的, uncreative无创造力的

mydata$extra.actor\_rated = (mydata$extra\_1.a + (8-(mydata$extra\_2.a)))/2

mydata$agree.actor\_rated = (mydata$agree\_2.a + (8-(mydata$agree\_1.a)))/2

mydata$conc.actor\_rated = (mydata$conc\_1.a + (8-(mydata$conc\_2.a)))/2

mydata$neuro.actor\_rated = (mydata$neuro\_1.a + (8-(mydata$neuro\_2.a)))/2

mydata$open.actor\_rated = (mydata$open\_1.a + (8-(mydata$open\_2.a)))/2

mydata$extra.observer\_rated = (mydata$extra\_1.o + (8-(mydata$extra\_2.o)))/2

mydata$agree.observer\_rated = (mydata$agree\_2.o + (8-(mydata$agree\_1.o)))/2

mydata$conc.observer\_rated = (mydata$conc\_1.o + (8-(mydata$conc\_2.o)))/2

mydata$neuro.observer\_rated = (mydata$neuro\_1.o + (8-(mydata$neuro\_2.o)))/2

mydata$open.observer\_rated = (mydata$open\_1.o + (8-(mydata$open\_2.o)))/2

mydata$obs\_extra = (mydata$obs\_extra\_1.a + (8-(mydata$obs\_extra\_2.a)))/2

mydata$obs\_agree = (mydata$obs\_agree\_2.a + (8-(mydata$obs\_agree\_1.a)))/2

mydata$obs\_conc = (mydata$obs\_conc\_1.a + (8-(mydata$obs\_conc\_2.a)))/2

mydata$obs\_neuro = (mydata$obs\_neuro\_1.a + (8-(mydata$obs\_neuro\_2.a)))/2

mydata$obs\_open = (mydata$obs\_open\_1.a + (8-(mydata$obs\_open\_2.a)))/2

# Trait Authenticity 特质真实性

#Q1：在大多数情况下我都忠于自己

#Q2：我是一个真实的人

#Q3：我在互动中真诚比技巧更多

#Q4：我可以更真实一些

mydata$perauth\_4.a.r = 8 - mydata$perauth\_4.a

mydata$perauth.actor\_rated = (mydata$perauth\_1.a + mydata$perauth\_2.a + mydata$perauth\_3.a + mydata$perauth\_4.a.r)/4

mydata$perauth\_4.o.r = 8 - mydata$perauth\_4.o

mydata$perauth.observer\_rated = (mydata$perauth\_1.o + mydata$perauth\_2.o + mydata$perauth\_3.o + mydata$perauth\_4.o.r)/4

mydata$obs\_perauth\_4.a.r = 8 - mydata$obs\_perauth\_4.a

mydata$obs\_perauth = (mydata$obs\_perauth\_1.a + mydata$obs\_perauth\_2.a + mydata$obs\_perauth\_3.a + mydata$obs\_perauth\_4.a.r)/4

对人口学变量重新赋值

#性别

mydata$man[mydata$Gender == 1] = 0

mydata$woman[mydata$Gender == 2] = 1

mydata$obs\_woman[mydata$obs\_Gender == 1] = 0

mydata$obs\_woman[mydata$obs\_Gender == 2] = 1

#人种

mydata$white[mydata$race\_ethnicity == 3] = 1

mydata$white[mydata$race\_ethnicity != 3] = 0

mydata$obs\_white[mydata$obs\_race\_ethnicity == 3] = 1

mydata$obs\_white[mydata$obs\_race\_ethnicity != 3] = 0

计算真实性差异

# Authenticity Accuracy Score

# Signed diff between perceiver rated authenticity and target-rated authenticity

#真实性差异=我对他人评价的真实性-自我评价真实性

mydata$auth\_diff = (mydata$perauth.observer\_rated - mydata$perauth.actor\_rated)

数据分析

首先对变量进行描述性统计：平均数和标准差

table(actor.summary$woman)

mean(actor.summary$age,na.rm = T)

sd(actor.summary$age,na.rm = T)

接下来计算统计检验力，看自我评价特质真实性是否显著

actor.percauth= data.frame(scale.items.summary$perauth\_1.a,scale.items.summary$perauth\_2.a,scale.items.summary$perauth\_3.a,scale.items.summary$perauth\_4.a.r)

psych::alpha(actor.percauth,check.keys = T)

remove(actor.percauth)

进行额外的真实性测量，计算相关

cor.test(scale.items.summary$cn1.r,scale.items.summary$cn2.r)

cor.test(scale.items.summary$cn1.2.r,scale.items.summary$cn2.2.r)

行为真实性的计算

acting.auth= data.frame(scale.items.summary$acting.authentic\_1.r,scale.items.summary$acting.authentic\_2,scale.items.summary$acting.authentic\_3)

psych::alpha(acting.auth,check.keys = T)

remove(acting.auth)

两个被了解的感觉的总分进行相关

cor.test(scale.items.summary$known\_1,scale.items.summary$known\_2)

计算统计检验力，看对他人评价特质真实性是否显著

obs.percauth= data.frame(scale.items.summary$perauth\_1.o,scale.items.summary$perauth\_2.o,scale.items.summary$perauth\_3.o,scale.items.summary$perauth\_4.o.r)

psych::alpha(obs.percauth)

remove(obs.percauth)

对真实性差异分数进行描述性统计

mean(mydata$auth\_diff,na.rm = T)

sd(mydata$auth\_diff,na.rm = T)

做相关性表格

Table1包含M,SD,相关系数和置信区间，并生成文档表格。

table= data.frame(actor.summary$perauth.observer\_rated,actor.summary$perauth.actor\_rated,actor.summary$fas.1,actor.summary$fas.2,actor.summary$acting.auth,actor.summary$known,actor.summary$open.a,actor.summary$conc.a,actor.summary$extra.a,actor.summary$agree.a,actor.summary$neuro.a,actor.summary$fam.o,actor.summary$age,actor.summary$woman,actor.summary$white)

apa.cor.table(table,filename = "study2\_correlation.doc")

大五人格特质的表格形成table2并生成文档表格。

table2= data.frame(actor.summary$extra.a,actor.summary$extra.o,actor.summary$open.a,actor.summary$open.o,actor.summary$conc.a,actor.summary$conc.o,actor.summary$agree.a,actor.summary$agree.o,actor.summary$neuro.a,actor.summary$neuro.o)

apa.cor.table(table2, filename = "study2\_supplementarytable.doc")

对相关表格进行作图，共四个图：真实性、状态真实性、行为真实性和被了解的感觉与他人评价真实性之间的相关关系图，生成p1、p2、p3、p4四个图的PDF形式。

p1 <- ggplot(mydata, aes(x=perauth.actor\_rated, y=perauth.observer\_rated)) +

geom\_jitter(color = "firebrick2", alpha = .5)+

geom\_smooth(method=lm, color="firebrick4")+

labs(x="Authenticity", y = "Other-Rated Authenticity")+

theme\_classic()

p2 <- ggplot(mydata, aes(x=fas.2, y=perauth.observer\_rated)) +

geom\_jitter(color = "green3", alpha = .5)+

geom\_smooth(method=lm, color="green4")+

labs(x="State Authenticity", y = "")+

theme\_classic()

p3 <- ggplot(mydata, aes(x=acting.auth, y=perauth.observer\_rated)) +

geom\_jitter(color = "darkorange2",alpha = .5)+

geom\_smooth(method=lm, color="darkorange4")+

labs(x="Acting Authentically", y = "")+

theme\_classic()

p4 <- ggplot(mydata, aes(x=known, y=perauth.observer\_rated)) +

geom\_jitter(color = "deepskyblue2",alpha = .5)+

geom\_smooth(method=lm, color="deepskyblue4")+

labs(x="Feeling Known", y = "")+

theme\_classic()

p1 + p2 + p3 + p4

#ggsave("selfvother.pdf")

进行假设检验

mydata[,c("perauth.observer\_rated","perauth.actor\_rated","fas.1","fas.2","known","acting.auth","extra.actor\_rated","conc.actor\_rated","agree.actor\_rated","neuro.actor\_rated","open.actor\_rated","extra.observer\_rated","conc.observer\_rated","agree.observer\_rated","neuro.observer\_rated","open.observer\_rated","fam.o")] <- apply(mydata[,c("perauth.observer\_rated","perauth.actor\_rated","fas.1","fas.2","known","acting.auth","extra.actor\_rated","conc.actor\_rated","agree.actor\_rated","neuro.actor\_rated","open.actor\_rated","extra.observer\_rated","conc.observer\_rated","agree.observer\_rated","neuro.observer\_rated","open.observer\_rated","fam.o")], 2, function(x) {as.numeric(as.character(scale(x, center = T, scale = T)))})

mydata$obsemail = as.factor(mydata$obsemail)

mydata$email = as.factor(mydata$email)

## Do others know when you're authentic? #你诚实的时候他人知道吗？

用多级估量法MLM分析多层次数据

summary(m1 <- lmer(perauth.observer\_rated ~ perauth.actor\_rated + (1 | obsemail), mydata, REML = FALSE))

confint(m1)

summary(m2 <- lmer(perauth.observer\_rated ~ fas.1 + (1 | obsemail), mydata, REML = FALSE))

confint(m2)

summary(m3 <- lmer(perauth.observer\_rated ~ fas.2 + (1 | obsemail), mydata, REML = FALSE))

confint(m3)

summary(m4 <- lmer(perauth.observer\_rated ~ acting.auth + (1 | obsemail), mydata, REML = FALSE))

confint(m4)

summary(m5 <- lmer(perauth.observer\_rated ~ known + (1 | obsemail), mydata, REML = FALSE))

confint(m5)

进行回归分析

# OLS

summary(m1 <- lm(perauth.observer\_rated ~ perauth.actor\_rated, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ fas.1, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ fas.2, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ acting.auth, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ known, mydata))

confint(m1)

进行贝叶斯分析

m1 = stan\_glm(perauth.observer\_rated ~ perauth.actor\_rated, data = mydata)

计算系数，信度和贝叶斯

coef(m1)

posterior\_interval(m1, prob = 0.95)

rope(m1, ci = .90, range = rope\_range(m1))

## SRM

smaller = read.csv("srm.csv")

## Interpreting SRM notes:

## Actor Variance / Effect Reliability == Perceiver/Observer; one making the rating

## Partner Variance / Effect Reliability == Target; one receiving the rating

## Relationship Variance == Observer+Target Variance

## Self rating with Actor == Do people who describe themselves as authentic see others as authentic?

## Self rating with Partner == Are people who describe themselves as authentic seen by others as authentic?

RR1 <- RR(perauth.observer\_rated~perceiver.id\*target.id | group.id, data = smaller, na.rm = T)

RR1

```

人格判断的准确性分析

summary(m1 <- lmer(extra.observer\_rated ~ extra.actor\_rated + (1 | obsemail), mydata, REML = FALSE))

confint(m1)

summary(m2 <- lmer(conc.observer\_rated ~ conc.actor\_rated + (1 | obsemail), mydata, REML = FALSE))

confint(m2)

summary(m3 <- lmer(agree.observer\_rated ~ agree.actor\_rated + (1 | obsemail), mydata, REML = FALSE))

confint(m3)

summary(m4 <- lmer(neuro.observer\_rated ~ neuro.actor\_rated + (1 | obsemail), mydata, REML = FALSE))

confint(m4)

summary(m5 <- lmer(open.observer\_rated ~ open.actor\_rated + (1 | obsemail), mydata, REML = FALSE))

confint(m5)

进行回归分析

# OLS

summary(m1 <- lm(extra.observer\_rated ~ extra.actor\_rated, mydata))

confint(m1)

summary(m1 <- lm(conc.observer\_rated ~ conc.actor\_rated, mydata))

confint(m1)

summary(m1 <- lm(agree.observer\_rated ~ agree.actor\_rated, mydata))

confint(m1)

summary(m1 <- lm(neuro.observer\_rated ~ neuro.actor\_rated, mydata))

confint(m1)

summary(m1 <- lm(open.observer\_rated ~ open.actor\_rated, mydata))

confint(m1)

贝叶斯分析

# Are Observers Overly Positive? #自己对他人的评价总是更积极，进行t检验。

t.test(actor.summary$perauth.observer\_rated,actor.summary$perauth.actor\_rated)

## Does Perceiver's Own Authenticity Predict Their Ratings of Authenticity?“感知者自己的真实性是否预测了他们的真实性评级？”这个问题

对观察者的变量进行标准化

mydata[,c("obs\_perauth","obs\_fas.1","obs\_fas.2","obs\_known","obs\_acting.auth","obs\_open","obs\_conc","obs\_extra","obs\_agree","obs\_neuro")] <- apply(mydata[,c("obs\_perauth","obs\_fas.1","obs\_fas.2","obs\_known","obs\_acting.auth","obs\_open","obs\_conc","obs\_extra","obs\_agree","obs\_neuro")],2,function(x) {as.numeric(as.character(scale(x, center = T, scale = T)))})

summary(m1 <- lmer(perauth.observer\_rated ~ obs\_perauth + (1 | email), mydata, REML = FALSE))

confint(m1)

summary(m2 <- lmer(perauth.observer\_rated ~ obs\_perauth + obs\_open + obs\_conc + obs\_extra + obs\_agree + obs\_neuro + fam.o + obs\_woman + obs\_white + obs\_Age + fam.o + (1 | email), mydata, REML = FALSE))

confint(m2)

summary(m3 <- lmer(perauth.observer\_rated ~ obs\_fas.1 + (1 | email), mydata, REML = FALSE))

confint(m3)

summary(m4 <- lmer(perauth.observer\_rated ~ obs\_fas.2 + (1 | email), mydata, REML = FALSE))

confint(m4)

summary(m5 <- lmer(perauth.observer\_rated ~ obs\_acting.auth + (1 | email), mydata, REML = FALSE))

confint(m5)

summary(m6 <- lmer(perauth.observer\_rated ~ obs\_known + (1 | email), mydata, REML = FALSE))

confint(m6)

summary(m7 <- lmer(auth\_diff ~ obs\_perauth + (1 | email), mydata, REML = FALSE))

confint(m7)

# OLS

进行回归分析

summary(m1 <- lm(perauth.observer\_rated ~ obs\_perauth, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ obs\_perauth + obs\_open + obs\_conc + obs\_extra + obs\_agree + obs\_neuro + fam.o + obs\_woman + obs\_white + obs\_Age + fam.o, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ obs\_fas.1, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ obs\_fas.2, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ obs\_acting.auth, mydata))

confint(m1)

summary(m1 <- lm(perauth.observer\_rated ~ obs\_known, mydata))

confint(m1)

summary(m1 <- lm(auth\_diff ~ obs\_perauth, mydata))

confint(m1)

# Who is perceived as authentic? “谁被认为是真实的？”对于这个问题进行探索

summary(m1 <- lmer(perauth.observer\_rated ~ extra.actor\_rated + conc.actor\_rated + agree.actor\_rated + open.actor\_rated + neuro.actor\_rated + woman + white + Age + (1 | obsemail), mydata, REML = FALSE))

confint(m1)

# OLS

进行回归分析

summary(m1 <- lm(perauth.observer\_rated ~ extra.actor\_rated + conc.actor\_rated + agree.actor\_rated + open.actor\_rated + neuro.actor\_rated + woman + white, mydata))

confint(m1)

```

额外的作标准差的图

vars <- c("obs\_fas.1", "obs\_fas.2", "obs\_acting.auth", "obs\_known","obs\_perauth")

coefs\_se\_model <- data.frame()

p = 1

for (i in vars) {model\_fit <- summary(lmer(perauth.observer\_rated ~ get(i) + (1 | email), mydata, REML = FALSE))

coefs\_se\_model[p, 1] <- i

coefs\_se\_model[p, 2] <- model\_fit$coefficients[2,1]

coefs\_se\_model[p, 3] <- model\_fit$coefficients[2,2]

coefs\_se\_model[p, 4] <- model\_fit$coefficients[2,3]

coefs\_se\_model[p, 5] <- model\_fit$coefficients[2,4]

p = p+1

}

coefs\_se\_model$Predictor <- as.factor(coefs\_se\_model$V1)

coefs\_se\_model$Predictor<-factor(coefs\_se\_model$Predictor,levels= coefs\_se\_model$Predictor[order(coefs\_se\_model$V2)])

coefs\_se\_model = coefs\_se\_model[order(coefs\_se\_model$V2),]

ggplot(coefs\_se\_model, aes(x=Predictor, y=V2, color=Predictor)) +

geom\_errorbar(aes(ymin=V2-V3, ymax=V2+V3), width=.1, position=position\_dodge(0.1)) +

geom\_point(position=position\_dodge(0.1), size =5) +

scale\_fill\_brewer(palette="Set1") +

xlab("Predictor") +

ylab("Regression Coefficient") +

theme\_classic()+

geom\_hline(yintercept=0, linetype="dashed", color = "black")+

coord\_flip() +

theme(axis.text.y = element\_blank())

对熟悉程度和准确性进行讨论

summary(m1 <- lm(perauth.observer\_rated ~ fam.o, mydata))

confint(m1)

1. 复现结果

该研究使用普通最小二乘(OLS)回归模型，在“真实性”“感觉真实性”“行为真实性”“被了解的感觉”四个项目上，比较了自我评定的特质真实性和他人报告的特质真实性的差异。首先，复现发现自我评价的特质真实性不能预测他人评价的特质真实性*(β* = 0.04, *SE* = 0.03, 95% *CI* = [- 0.01, 0.09]，*p* = .132)。其次，在第一次和第四次调查中，由他人评定的特质真实性与自我评定的状态真实性也无关(调查1: *β* = 0.01, *SE* = 0.03, 95% *CI* = [- 0.04, 0.06]，*p* = .608;调查4: *β* = 0.02, *SE* = 0.03, 95% *CI* = [-0.03, 0.08]，*p* = .348)。接着，由他人评价的特质真实性也与被试对自己能够与队友真诚合作的感知无关(*β* = 0.00, *SE* = 0.03, 95% *CI* = [- 0.05, 0.05]，*p* = .980)。最后，由他人评价的特质真实性也无法预测被试对队友认识真实的他们这一维度上的评分(*β* = - 0.04, *SE* = 0.03, 95% *CI* = [- 0.09, 0.01]，*p* = .142)。综上所述，四种结果共同表明他人评价的真实性与自我评价的真实性不相关，与文献结果一致，具体复现结果见图1。

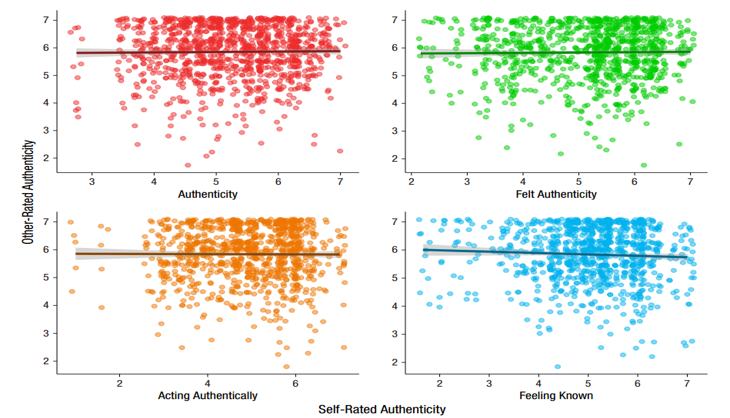


图1四个项目自我评分与他人评分的OSL回归模型

为了是数据更可靠，复现使用非信息先验估计了自我评价和他人评价真实性之间关系的贝叶斯回归，以获得回归系数的可信区间(Wagenmakers et al，2016)。研究发现，他人评定特质真实性系数对自我评价特质真实性系数的95%可信区间为[-0.05,0.07]，包含0。此外，我们使用实际等效区域(ROPE)方法测试了两者的无相关关系(Kruschke & Liddell, 2018)。研究发现，在自我评价和他人评价的特质真实性之间的关系上，HDI和ROPE完全重叠(100.00%)。这表明零假设没有被拒绝，即自我评价真实性与他人评价的真实性之间没有显著的相关关系，与文献结果一致。

为了检验关系的哪个方面解释了真实性评级的差异，复现与文献结果一致，进行了SRM(Social-relations modeling) (Kenny & La Voie, 1984)分析，它将自我和他人评级的方差分解为：做出判断的人、导致判断的目标特征以及自我-他人配对特有的这三个部分。此外，SRM在控制群体成员关系的条件下，还提供了他人评定与自我评定的部分相关性(即，真实性与被认为真实之间的相关性是什么?)，以及被试评分与他们给出的他人评价的相关性。

复现SRM的结果表明，目标间的差异，解释了约6%的方差(*SE* = 0.02, *p* = 0.003)。另一方面，被试间差异解释了真实性评分中41%的方差(*SE* = 0.06, *p* < .001)，自我-他人关系的差异解释了真实性评分中45%的方差(*SE* = 0.03, *p* < .001)。这表明，真实性评级的大部分差异与做出评分的被试之间的差异以及自我-他人不同关系之间的差异有关。

在控制群体成员的情况下，自评特质真实性与他人评价特质真实性的偏相关不显著，*r* = 0.12, *t*(167) = 1.53, *p* = 0.129。然而，在控制了群体成员关后，评价者的真实性与他们对他人真实性的感知之间存在显著的偏相关，*r* = 0.34, *t*(167) = 4.69, *p* < 0.001，与文献结果一致。

为了这个样本中的评分者并不能准确地判断他们的目标这一可能性，复现与文献结果一致，对自评人格特质和他评人格特质再次进行了一系列多层次模型，两者的正相关关系表明，自评人格特质得分对他评人格特质得分具有正向预测作用，这一结果表明评价者的人格评价是准确的。

复现与文献结果一致发现，他人评定的特质真实性(*M* = 5.84)显著高于自我评定的特质真实性(*M* = 5.27;*SD*= 0.57,95% *CI* = [0.43, 0.71])，*t*(315) = 7.98, *p* < 0.001, *Cohen’s d* = 0.79。也就是说，参与者更有可能认为其他人比他们自己更真实。

同时，复现与文献结果一致，发现被试的特质真实性与其对其他个体的真实性评分之间存在显著的正相关关系(*β* = 0.23, *SE* = 0.03, 95% *CI* = [0.17, 0.29]，p < .001)。当控制了评分者的个性、人口统计学特征和对目标的熟悉程度时，这种效应仍然存在(*β* = 0.20, *SE* = 0.03, 95% *CI* = [0.13, 0.26]，*p* < .001)。

为了解释评分者的偏见是由于项目的相似性，而不是他们自己的真实性这一观点，研究控制了其他真实性变量(状态、行为真实、被了解的感觉)与真实性评级之间的关系。复现与文献结果一致发现调查1和调查4这两个时间点的状态真实性对其他个体的真实性判断有显著的正向影响(调查1: *β*= 0.12, *SE* = 0.03, 95% *CI* = [0.06, 0.18]，*p* < .001;调查4: *β*= 0.11, *SE* = 0.03, 95% *CI* = [0.04, 0.17]，*p* < .001)。

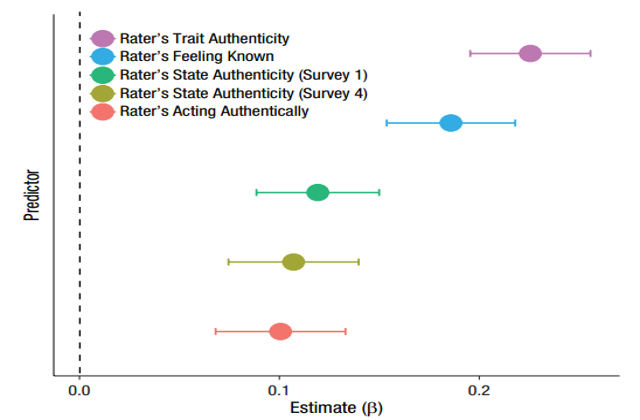


图2 五种自我真实性指标预测对他人真实性判断

复现与文献结果一致，还发现被试认为自己的行为是真实的可以正向预测被试对他人的真实性判断(*β*= 0.10, *SE* = 0.03, 95% *CI* = [0.04, 0.17]，*p* = 0.002)，同样也可以正向预测被试对群体成员被了解的感受的真实性判断(*β* = 0.19, *SE* = 0.03, 95% *CI* = [0.12, 0.25]，*p* < 0.001) 。然后，研究以被试的真实性作为符号差异真实性-错误得分的预测因子。被试的特质真实性预测了正的真实性-错误评分(*β* = 0.24, *SE* = 0.03, 95% *CI* = [0.18, 0.30]，*p* < .001)。也就是说，真实的被试认为他评真实性相较于自评真实性更真实。

最后，研究为了探索那些被试更有可能被评为真实的，对目标的人口统计学和人格变量进行了探索性分析，并在评分水平上聚类，。结果没有发现存在任何个体差异可以预测他人特质真实性的评价 (*ps* > .11)。

4讨论

4.1真实性判断的准确性

总体来说，他人评价的真实性与自我评价的真实性不相关。具体表现在三个方面：在特质真实性方面，发现自我评价的状态真实性不能预测他人评价的状态真实性；在状态真实性方面，在第一次和最后一次调查中，他人评定的特质真实性与自我评定的状态真实性无关；在行为真实性和被了解的感觉方面，他人评价的特质真实性与被试对自己能够与队友真诚合作的感知无关，他人评价的特质真实性不能通过参与者的评价来预测他们的队友认识他们。

评分的差异。评价者之间的差异解释了真实性评分中41%的方差，自我-他人关系的差异解释了真实性评分中45%的方差。这表明，真实性评级的大部分差异与做出评级的人之间的差异以及不同关系之间的差异有关，而不是因为目标的差异。

4.2感知的真实性

总体来说，他人评价的特质真实性高于自我评价的特质真实性，人们更认为其他人比他们自己更真实，这说明感知真实性是存在偏见的。例如，个体认为他们的积极行为比消极行为更真实(Jongman-Sereno & Leary, 2016)，积极情绪比特质一致的表达更能预测自我评价的真实性(Lenton et al, 2013)。如果自评真实性受到情境或效价特征的影响，它会增加其可感知性的额外压力。

评价者的特质真实性。发现评分者的特质真实性与其对其他个体的真实性评分之间存在显著的正相关关系。评级者的状态真实性。发现两个时间点的状态真实性对其他个体的真实性判断有显著的正向影响。评分者的行为是真实的，被了解的感觉是存在的。同样，评分者认为他们可以真实地行事正向预测其对他人的真实性判断，相对于被试对自己真实性的评价，真实评分者认为其他人更真实。

最后发现没有个体差异预测其他评价的特质真实性，说明元知觉（期望别人认为你是真实的）同样与他人对真实性的评价无关。这就说明人们很难准确的感知他人的知识性。可以解释为：首先，真实性可能很难观察，而非真实性则不然。对欺骗准确性的研究发现了一个类似的模式——人们更准确地识别谎言是虚假的，而不是识别真理是真实的(Bond & DePaulo, 2006)。其次，伪造真实性的特征可能比真实地满足它们更容易。以装病为例:假装生病的人比真正生病的人更有可能表现出疾病的所有症状，而真正生病的人只表现出一些症状，而没有其他症状。

4.3研究局限

首先，没有观察到这些真实性判断的有用性。或许那些彼此真实的群体是否比不认为彼此真实的群体表现的更好，有更多的信任和凝聚力。

其次，研究人群主要集中在土耳其工人和MBA学生当中，研究结果的普遍性难以保证。

最后，更多的是选择在特质层面去比较真实性评级，但忽视了状态层面的测量，没有将感知真实性和真实性的多方面结构充分的结合起来。

该研究也没有考虑到感知真实性偏差发生的时间和原因，因为个人进行真实性的自我评价时他们是谦虚的，在进行这些评级时可用信息存在不对称性，需要进一步的经验证据更深入地理解进行这种偏见。

该研究还发现熟悉度没有增加感知真实性的准确性，可能是由于许多动机造成的，未来可以进一步考虑熟悉度与感知真实性之间的关系。

本文的核心贡献是对于真实性的研究，推动了真实性前因的研究，将重点放在个人特征上并与自我评价的真实性有关。但也有例外，感知的真实性可能与社会规范有更多的关系。

5 参考文献

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256–274.

Bailey, E. R., & Levy, A. (2022). Are You for Real? Perceptions of Authenticity Are Systematically Biased and Not Accurate. *Psychological Science*, *33*(5),798–815.

Ehrhart, M. G., Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences*, *47*(8), 900–905.

Kenny, D. A., & La Voie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 141–182). Academic Press.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psycho-nomic Bulletin & Review*, *25*(1), 178–206.

Wagenmakers, E.-J., Morey, R. D., & Lee, M. D. (2016). Bayesian benefits for the pragmatic researcher. *Current Directions in Psychological Science*, *25*(3), 169–176.

Wood, A. M., Linley, P. A., Maltby, J., Baliousis, M., & Joseph, S. (2008). The authentic personality: A theoretical and mpirical conceptualization and the development of the Authenticity Scale. *Journal of Counseling Psychology*, *55*(3), 385–399.

6 分工

R代码复现：艾骄阳、孙玉璇、王健、张雨婷；

文档撰写及反馈修改：艾骄阳、王健；

PPT制作：张雨婷；

汇报：孙玉璇。

To the words " Are You for Real? Perceptions of Authenticity Are Systematically Biased and Not Accurate " reproducibility study

**Abstract** Can people accurately perceive that others are real? The goal of this study is to compare the authenticity of self-evaluation and other evaluations among randomly assigned subjects. The hypothesis is that the relationship between self-evaluation and the authenticity of others' evaluation is not significant, and the results also verify the hypothesis. In addition, the study also tested for a bias in perceived authenticity. Firstly, the authenticity of others 'evaluation and self-evaluation are compared, and the results find that the authenticity of others' evaluation is higher than the authenticity of self-evaluation. It was also found that the rater's own authenticity could predict their rating of the authenticity of others. Overall, we found that people cannot accurately tell whether others are real.

**Keywords** self-evaluation, others' evaluation, authenticity