The True Role that Suppressor Effects Play in Condition-Based Regression Analysis: None. A Reply to Fiedler (2021)

Supplemental Online Material

Contents

1 Mathematical Proof of the Suppressor Effect

Fiedler uses a simulation study to demonstrate that the regression weight c_2 in the regression model $H = c_0 + c_1 S + c_2 R + \varepsilon$ can be negative when the correlation r_{RH} between the reality criterion R and happiness H is zero (pages 3.5 – 7 of Fiedler's comment). In Section 1.1, we provide the formal proof for this mathematical fact. In Section 1.2, we provide an example computation that can help readers understand the proof. In Section 1.3, we show how one can generate data with arbitrary choices for the correlations between S, R, and H, that can then be used to *estimate* the model coefficients in a random sample. We also show how the latter code can be used to reproduce Fiedler's simulation study.

Note that, as we explain in our paper, suppressor effects and thus the results of the proof and the simulation study are unrelated to CRA. Nonetheless, we provide this information because we consider it important for readers to be able to completely reproduce the results in the discourse in order to advance their understanding of the properties of CRA. Readers might also want to use this code to play around with the distributions of the variables (e.g., chose arbitrary correlations, means, and/or variances) to fully understand the attributes of the data and the corresponding results of CRA. Please see the file "ComF_SOM_Rcode.R" (provided in the OSF) for the code to reproduce the contents of this document.

1.1 Proof

In general, the coefficient c_2 in the regression model $H = c_0 + c_1 S + c_2 R + \varepsilon$ is given by the following formula (with SD = standard deviation; see, e.g., Fox, 2016):

$$c_2 = \frac{r_{RH} - r_{SH} \cdot r_{SR}}{1 - r_{SR}^2} \cdot \frac{SD(H)}{SD(R)} \tag{1}$$

Now let $r_{RH} = 0$, $r_{SH} > 0$, and $r_{SR} > 0$ (as in Fiedler's simulation), then the following inequalities hold:

- $r_{SH} \cdot r_{SR} > 0$ (because both correlations are positive),
- $r_{RH} - r_{SH} \cdot r_{SR} = 0 - r_{SH} \cdot r_{SR} < 0$ (because $r_{SH} \cdot r_{SR} > 0$),
- $1 - r_{SR}^2 > 0$ (because $r_{SR}^2 \in [0, 1]$),
- and $\frac{SD(H)}{SD(R)} > 0$ (because both standard deviations are positive).

When put together in Equation 1, these facts imply that $c_2 < 0$.

1.2 Computational Examples

To trace back the proof with an example, let S, R, and H be standard normal variables (which in particular implies $SD(H) = SD(R) = 1$) with correlations of $r_{RH} = 0$, $r_{SH} = 0.4$, and $r_{SR} = 0.4$. These example correlations are a special case of the correlations considered in the proof. Likewise, the inequalities that were considered above hold:

- $r_{SH} \cdot r_{SR} \; = \; 0.4 \cdot 0.4 \; = \; 0.16 \; > \; 0,$
- $r_{RH} - r_{SH} \cdot r_{SR} \; = \; 0 - 0.16 \; = \; -0.16 \; < \; 0,$
- $1 - r_{SR}^2 \; = \; 1 - 0.4^2 \; = \; 1 - 0.16 \; = \; 0.84 \; > \; 0,$
- and $\frac{SD(H)}{SD(R)} \; = \; \frac{1}{1} \; = \; 1 > 0.$

When put together in Equation 2, these results imply that:

$$\begin{aligned} c_2 &= \frac{r_{RH} - r_{SH} \cdot r_{SR}}{1 - r_{SR}^2} \cdot \frac{SD(H)}{SD(R)} \\ &= \frac{-0.16}{0.84} \cdot 1 \\ &= -0.1904762 \; < \; 0 \end{aligned} \tag{2}$$

The following table shows the standardized regression coefficients c_2 and also c_1 that could likewise be computed for all combinations of correlations reported for Fiedler's simulation. Here, the formula used to compute c_1 parallels the one for c_2 (e.g., Fox, 2016):

$$c_1 = \frac{r_{SH} - r_{RH} \cdot r_{SR}}{1 - r_{SR}^2} \cdot \frac{SD(H)}{SD(S)} \tag{3}$$

Table 1
Standardized Coefficients

rSR	rSH	c1	c2
0.2	0.2	0.208	-0.042
0.2	0.4	0.417	-0.083
0.2	0.6	0.625	-0.125
0.2	0.8	0.833	-0.167
0.4	0.2	0.238	-0.095
0.4	0.4	0.476	-0.190
0.4	0.6	0.714	-0.286
0.4	0.8	0.952	-0.381
0.6	0.2	0.312	-0.188
0.6	0.4	0.625	-0.375
0.6	0.6	0.938	-0.562
0.6	0.8	1.250	-0.750
0.8	0.2	0.556	-0.444
0.8	0.4	1.111	-0.889
0.8	0.6	1.667	-1.333

[a] Note. rRH = 0 in all cases.

1.3 Simulated Examples (Including Code to Reproduce Fiedler's Simulation Study)

To further exemplify the formula and its consequences, we show how one can generate data with arbitrary choices for the correlations between S, R, and H, that can then be used to *estimate* the model coefficients in a random sample.

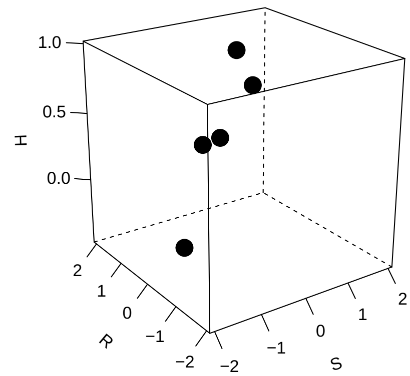
To prepare the simulation, we used the `mvtnorm` package (Genz & Bretz, 2009; Genz et al., 2020), which enables users to draw random data from a multivariate normal distribution. We then define a function `drawdat` that samples random variables S, R, and H for any choice of correlations r_{SR}, r_{SH}, and r_{RH} and returns a data frame with these three variables. The variables are sampled from a multivariate *standard* normal distribution by default (i.e., with means of zero and standard deviations of 1), but the function allows the user to choose arbitrary means and standard deviations for the variables.

```r
library(mvtnorm)

drawdat <- function(rSR, rSH, rRH,
                    n = 100,
                    mS = 0, mR = 0, mH = 0, # means
                    sdS = 1, sdR = 1, sdH = 1, # SDs
                    seed = 1909
                    ){
```

```
  # compute variances
  vS <- sdS^2
  vR <- sdR^2
  vH <- sdH^2

  # compute covariances
  covSR <- rSR*sdS*sdR
  covSH <- rSH*sdS*sdH
  covRH <- rRH*sdR*sdH

  # set a seed so that results will be reproducible
  set.seed(seed)

  # draw data from multivariate normal distribution
  dat <- rmvnorm(n, mean=c(mS, mR, mH),
               sigma=matrix(c(vS, covSR, covSH,
                              covSR, vR, covRH,
                              covSH, covRH, vH),
                            ncol=3))

  # prettify the data
  dat <- as.data.frame(dat)
  names(dat) <- c("S", "R", "H")

  # return the data
  return(dat)
}
```

For example:

```
# draw data
mydata <- drawdat(rSR = .4, rSH = .4, rRH = 0, n = 5)

# inspect the data
mydata

##             S         R          H
## 1  0.02824826 0.7083524  0.3094079
## 2 -0.74516241 0.8519169 -0.4919469
## 3 -0.07064318 0.0134388  1.0171224
## 4 -0.14630164 1.0620503  0.2282681
## 5  1.16717602 1.2220614  0.5647027

# plot the data
library(RSA)

plotRSA(points=list( data=mydata[,c("S","R","H")], show=TRUE, cex=3),
        suppress.surface=TRUE, param=FALSE, legend=FALSE)
```



The drawdat function can be used to reproduce Fiedler's simulation study. To prepare the simulation, we save the correlations that will be included in the simulation in respective lists. We also create an empty object (result) that will be used to store the results.

```
rSR.list <- c(.2, .3, .4, .5, .6, .7, .8, .9)
rSH.list <- c(0, .1, .2, .3, .4, .5, .6, .7, .8, .9, 1)
rRH.list <- c(0)

result <- NULL
```

The following code runs the simulation. It loops through the lists of prespecified correlations (rSR.list etc.). For each combination of correlations, data with the respective correlations are drawn with the drawdat function, which was defined above, and the linear regression model $H = c_0 + c_1 S + c_2 R + \varepsilon$ is estimated with the lm function. The estimated coefficients are then extracted and stored in the result object.

```
for(irSR in rSR.list) {

  for(irSH in rSH.list) {

    for(irRH in rRH.list) {

    # draw data
    data <- drawdat(rSR = irSR, rSH = irSH, rRH = irRH, n = 100)

    # standardize the variables (to obtain standardized coefficients)
    data <- data.frame(apply(data, 2, scale))

    # estimate the linear model
    fit <- lm(H ~ S + R, data=data)
```

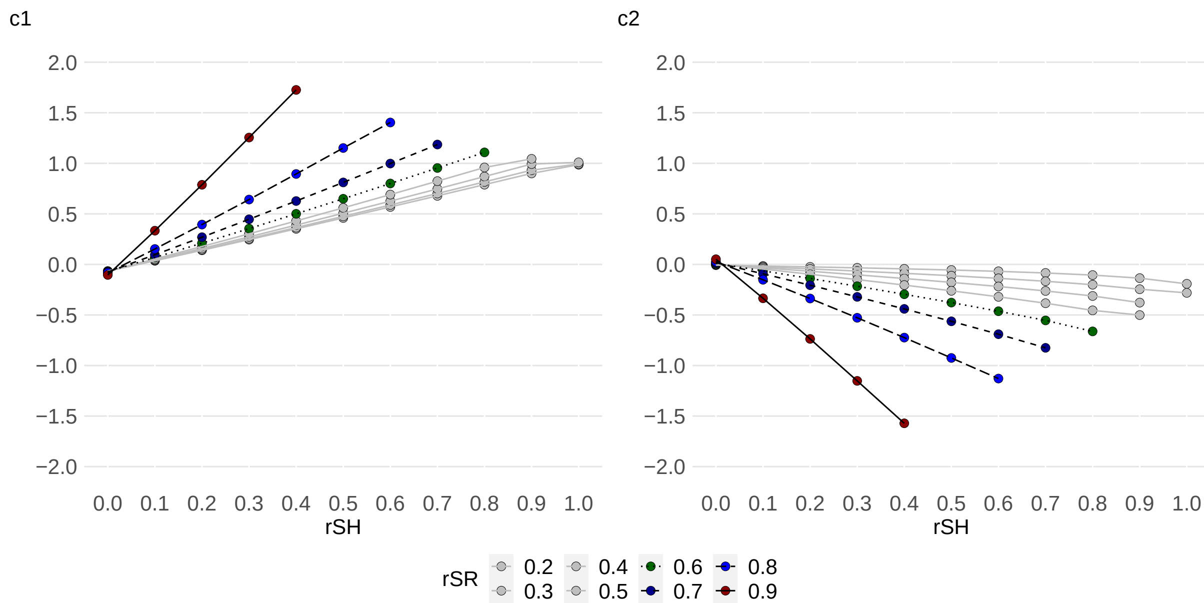
```
    # extract coefficients
    coefs <- summary(fit)$coefficients

    # store the results
    result <- rbind(result,

                    data.frame(
                      rSR = irSR,
                      rSH = irSH,
                      rRH = irRH,

                      c1 = coefs["S", "Estimate"],
                      c2 = coefs["R", "Estimate"]) )
    }
  }
}

# show results
result
```

The following table shows the results (only an excerpt of the output is shown for conciseness). Note that the estimates of the coefficients in the simulation can of course differ from the coefficients computed in Section 1.2 (and also from Fiedler's results) due to random sampling.

Table 2
Simulation Results

rSR	rSH	rRH	c1	c2
0.2	0.2	0	0.140	-0.025
0.2	0.4	0	0.352	-0.044
0.2	0.6	0	0.569	-0.068
0.2	0.8	0	0.789	-0.106
0.4	0.2	0	0.158	-0.068
0.4	0.4	0	0.389	-0.139
0.4	0.6	0	0.626	-0.218
0.4	0.8	0	0.869	-0.312
0.6	0.2	0	0.210	-0.139
0.6	0.4	0	0.500	-0.295
0.6	0.6	0	0.800	-0.463
0.6	0.8	0	1.108	-0.663
0.8	0.2	0	0.394	-0.337
0.8	0.4	0	0.894	-0.724
0.8	0.6	0	1.404	-1.129

The following figure provides a visual impression of the results and shows that the observed patterns are similar (up to sampling variance) to the results that Fiedler obtained in his simulation study (compare this plot with Fiedler's Figure 1).

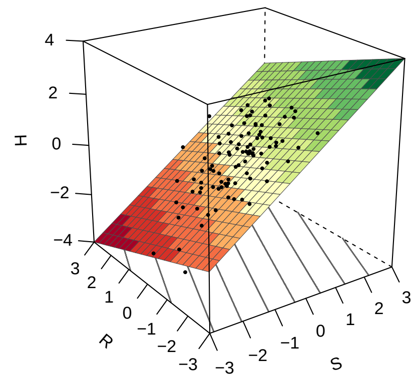


Please note again that suppressor effects, and thus the results of the simulation study, are unconnected to the question of whether an S-R effect pattern is present in the data. In contrast to Fiedler's claim, all data that were generated in the simulation study show an S-R effect pattern. This is accurately detected by CRA (given a sufficiently large sample size), as c_1 is positive and c_2 is negative for all of the data (see Table 2).

2 How to Inspect Example Data With Arbitrary PPZ-Correlation-Structures

The `drawdat` function defined above can be used, for example, to see that any data with a PPZ-correlation-structure (i.e., with $r_{SR} > 0, r_{SH} > 0, r_{RH} = 0$) show an S-R effect pattern that is accurately detected with CRA (given a large enough sample size so that the statistical tests are significant). The following code shows how to do this, for the example correlations $r_{SR} = .4, r_{SH} = .8, r_{RH} = 0$. It can be easily adapted to any choice of correlations (see the file "ComF_SOM_Rcode.R" for the plain R code).

```
# draw data for arbitrary choices of the correlations
# Note: set the sample size n to quite a large number to achieve that the
# correlations in the random sample will be close to those chosen for the
# population model.
mydata <- drawdat(rSR = .4, rSH = .8, rRH = 0, n = 100000)

# show the correlations in the random sample (as a check)
cor(mydata)

##           S            R            H
## S 1.0000000  0.396964261  0.800783187
## R 0.3969643  1.000000000 -0.001912714
## H 0.8007832 -0.001912714  1.000000000

# estimate the linear model
fit <- lm(H ~ S + R, data=mydata)

# inspect CRA results: inspect the coefficients
coefs <- summary(fit)$coefficients
coefs

##                 Estimate  Std. Error     t value  Pr(>|t|)
## (Intercept)  0.002487991 0.001540098    1.615476 0.1062109
## S            0.952620243 0.001680568  566.844185 0.0000000
## R           -0.381058718 0.001684935 -226.156386 0.0000000

# save coefficients to use them for the plot
c0 <- coefs["(Intercept)", "Estimate"]
c1 <- coefs["S", "Estimate"]
c2 <- coefs["R", "Estimate"]

# plot the estimated model (show only 100 data points);
# for rather flat surfaces, adapt the z axis limits: e.g., zlim=c(-2,2)
RSA::plotRSA(b0=c0, x=c1, y=c2,
             xlim=c(-3,3), ylim=c(-3,3), zlim=c(-4,4),
             points=list( data=mydata[1:100,c("S","R","H")], values="precicted" ),
             axes=c(), hull=F, param=F, legend=F)
```



As can be seen in the graph of the model, people with high values of $S - R$ (e.g., those in the green area of the regression surface) tend to have higher values of H than people with lower values of $S - R$ (e.g., those in the orange area) – the data show an S-R effect pattern. CRA detects this: $c_1 = 0.95$ is significantly positive and $c_2 = -0.38$ is significantly negative. That is, data with correlations of $r_{SR} = .4$, r_{SH}.8, and $r_{RH} = 0$, like any data with a PPZ-correlation-structure, show an S-R effect pattern that CRA accurately detects.

3 Suppressor Effects are Unconnected to SE Effect Patterns: Empirical Illustration

Bivariate correlations can, by definition of the phenomenon of interest, not inform about the presence or absence of an S-R effect pattern. Complementing the conceptual explanation that we provide in the article, this fact can also be illustrated with empirical examples in two ways. First, the example data in Figure 1 shows that data with (Figure 1a) or without (Figure 1b) a suppressor effect can show an S-R effect pattern of the same kind and strength. This illustrates the fact that the question of whether the data has a suppressor effect and the question of whether it shows an S-R effect pattern are completely unrelated. Second, the correlation $r_{RH} = 0$ can occur in data with (Figure 1a) or without (Figure 1c) an S-R effect pattern, which mirrors the fact that this correlation cannot inform about the existence of an S-R effect pattern.

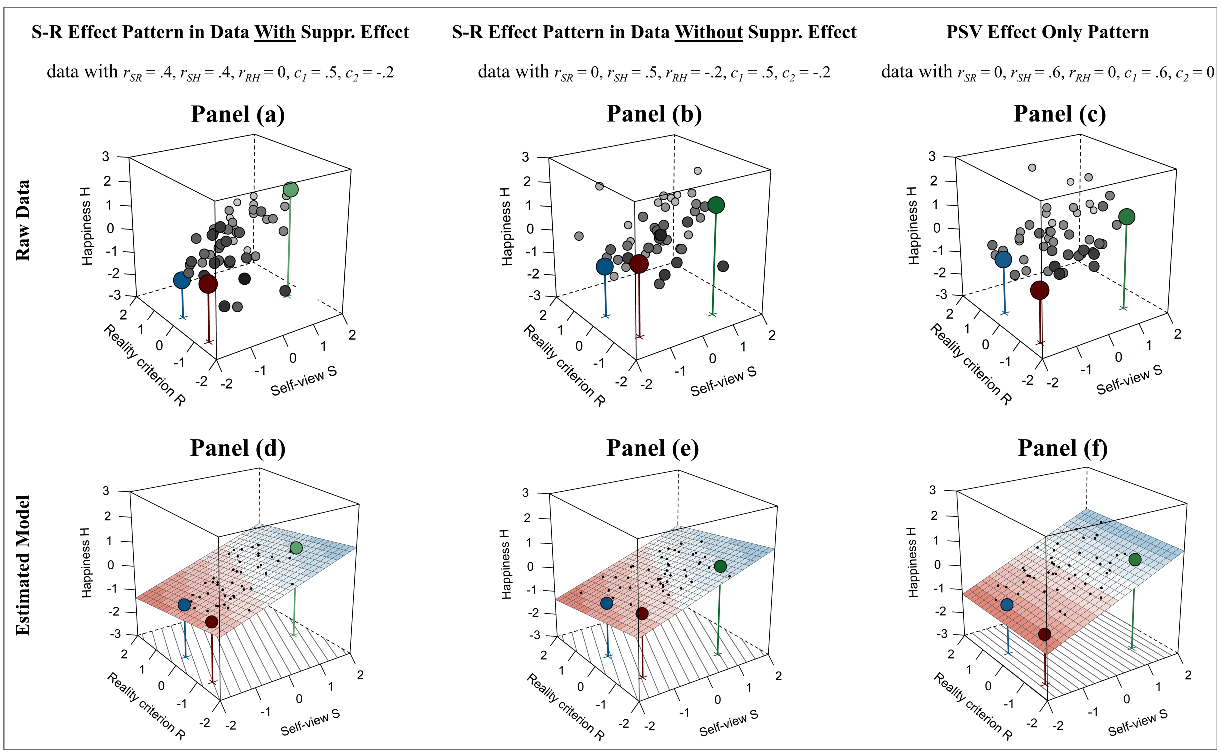


Figure 1. Example data and corresponding graphs of the fitted regression models (Equation 1 in the article)

4 Summary of Prior Discussions of Suppressor Effects in the Context of CRA

Fiedler's comment might evoke the impression that we ignored the possibility of suppressor effects when introducing CRA. This is not the case, but suppressor effects were discussed in both CRA articles and in our publicly available response to a review by Fiedler. In the following, we provide the specific references and summarize the relevant contents.

In the articles in which we introduced CRA (Humberg et al., 2018a; `http://doi.org/10.1037/pspp0000134`; see `https://www.osf.io/smmh7` for the openly accessible preprint), we discussed suppressor effects as follows (note that we used the term "SE effect" for an S-R effect pattern in that work):

> "A positive SE effect includes a negative effect of real ability in the regression with self-rated and real ability as predictors. It is important to note that this does *not* contradict the well-established positive effects of ability on a variety of beneficial outcomes . . . . It is possible that real ability is positively correlated with an outcome but has a negative relation with the very same outcome when self-viewed ability is controlled for, as is done in the regression that forms the basis of the CRA.[12] If, however, the effect of real ability on the outcome is still positive when controlling for self-rated ability, then there is definitely no *positive* SE effect because such an effect would require that out of two persons with the same self-viewed ability, the one who has less real ability (and not more, as would be the case in this situation) is happier."

Here, Footnote 12 reads:

> "Given that the correlation of real ability R and the outcome Z is positive, the coefficient c_2 is negative if and only if the relation $\mathrm{Cor}(R, Z) < \mathrm{Cor}(S, Z) \times \mathrm{Cor}(S, R)$ holds. This is the case, for example, if $\mathrm{Cor}(R, Z) = .1, \mathrm{Cor}(S, Z) = .4$ and $\mathrm{Cor}(S, R) = .3$."

Our second article on CRA (Humberg et al., 2018b; see `https://www.collabra.org/article/10.1525/collabra.137/` for the open access article) was a comment on an article by Krueger et al. (2017). Footnote 6 of our comment reads:

> "CRA focuses on the associations of S and T with H, respectively, when the respective other predictor is controlled for (i.e., semi-partial correlation). By contrast, the raw correlations $Cor(S, H)$ and $Cor(T, H)$ are not relevant to the existence of an SE effect and are thus not considered in the CRA approach. Given that semi-partial correlations can differ from raw

> correlations (e.g., see Horst, 1994; Smith, Ager, & Williams, 1992), this implies that there can be an SE effect even if, for example, the raw correlation between *T* and *H* is positive (see also Humberg et al., 2018)."

We included this footnote to address concerns about suppressor effects that Klaus Fiedler voiced as a reviewer of our comment. The review correspondence is openly accessible at `https://www.collabra.org/article/10.1525/collabra.137/`. In his review, Klaus Fiedler raised the following concerns about suppressor effects (see Reviewer 1 at `http://doi.org/10.1525/collabra.137.pr`):

> "My second point is much more concrete and it provides some explanation and illustration of what I have in mind in the preceding comment. The additive CRA model is based on the (unrealistic) assumption that S and T are uncorrelated (i.e., additive predictors). This is an assumption that I can't adopt. Why should S (self-rated ability) not be somehow correlated with T (actual ability)? Then, of course, the additive model is no longer warranted.
>
> One tricky implication is that, given two correlated predictors, r(S,T) > 0, the additive regression model will mimic the double condition, β1 > 0 and β2 < 0, even when H is exclusively dependent on S. Look what I have done to illustrate this point: I have started generating three (standardized) random variables S, T, and H (n = 100) under the following constraints:
>
> r(S,T) = .4
>
> r(S,H) = .6
>
> r(T,H) = 0; that is, H is unrelated to T; it is only related to S.
>
> An analysis of H regressed on S and T (as additive predictors) shows:
>
> βS +.71 and βT –.29 (both standard errors are roughly .084)
>
> Similar demonstrations of mimicry (of negative βT in spite of a zero causal impact) can be found with weaker correlations, although the negative βT gets smaller. However, this is not the point here. It is not my burden to show that the kind of mimicry I am referring to is always significant. It is the proponents' burden to show that the CRA is unbiased. Frankly speaking, it is hard for me to understand why the present debate seems to completely ignores the (counter-intuitive and hardly tenable) CRA assumption that S and T are fully independent, and of course also the implications of the alternative assumption that S and T may be substantially correlated."

We clarified these concerns as follows in our response letter (for the whole letter, see `https://s3-eu-west-1.amazonaws.com/ubiquity-partner-network/ucp/journal/collabra/137-2011-1-SP.docx`):

> "There appears to be a misunderstanding. CRA does not assume uncorrelated predictors, but its only assumptions are those of a multiple linear regression model (see also E.2). Thus, weak to moderate correlations among predictors (as are typical for ability or personality self-assessments) are unproblematic. When the predictors are highly correlated (i.e., over .9, see E.2), however, this can result in estimation problems in the multiple regression model. We now refer to this potential fallacy in the revised manuscript (see E.2).
>
> Note that the potential occurrence of mimicry (or suppressor effects, see R1.1) as another consequence of correlated predictor variables does not play a role for the CRA. The reason is that the CRA only involves the association of S and H with T controlled and the association of T and H with S controlled. Considering these controlled effects simultaneously allows to tell whether H relates to the difference S-T. The raw correlations, by contrast, are irrelevant for this question: The raw correlation of H and S, for example, includes no information about T and can thus not provide any information about the association of H and S-T. This also implies that the relation of the raw correlations and the regression weights (e.g., the question whether mimicry occurs) is not relevant for the question in the focus.
>
> As an example, consider the simulation study in the reviewer's comment: The reviewer's assumption that there is no SE effect in this data was based on the observation (in this case, construction) of the zero correlation of T and H. This raw correlation does, however, provide no information about the (non)existence of an SE effect. In fact, the CRA leads to a valid conclusion in the simulation: There is a positive association of S and H when T is controlled, which means that when comparing two people with equal values of ability T, the person who has the higher self-view and thus the higher value on the discrepancy (S-T) will be predicted to have a higher happiness H. There is a negative association of T and H when S is controlled, which means that for two people with equal self-views S, the person who has the lower ability level, who is again the person with a higher SE value (S-T), will be predicted to be happier."

5 References

Fox, J. (2016). *Applied regression analysis and generalized linear models.* Thousand Oaks, USA: SAGE Publications.

Genz, A., Bretz, F. (2009). *Computation of Multivariate Normal and t Probabilities*, series Lecture Notes in Statistics. Springer-Verlag, Heidelberg. ISBN 978-3-642-01688-2

Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., Hothorn, T. (2020). *mvtnorm: Multivariate Normal and t Distributions.* R package version 1.1-1. Retrieved from `https://CRAN.R-project.org/package=mvtnorm`.

Humberg, S., Dufner, M., Schönbrodt, F.D., Geukes, K., Hutteman, R., van Zalk, M.H.W., Denissen, J.J.A., Nestler, S., & Back, M.D. (2018a). Enhanced versus simply positive: A new condition-based regression analysis to disentangle effects of self-enhancement from effects of positivity of self-view. *Journal of Personality and Social Psychology, 114*(2), 303-322. `http://doi.org/10.1037/pspp0000134`

Humberg, S., Dufner, M., Schönbrodt, F.D., Geukes, K., Hutteman, R., van Zalk, M.H.W., Denissen, J.J.A., Nestler, S., & Back, M.D. (2018b). Why Condition-Based Regression Analysis (CRA) is Indeed a Valid Test of Self-Enhancement Effects: A Response to Krueger et al. (2017). *Collabra: Psychology, 4*(1), 26. `http://doi.org/10.1525/collabra.137`

Krueger, J. I., Heck, P. R., & Asendorpf, J. B. (2017). Self-enhancement: Conceptualization and Assessment. *Collabra: Psychology, 3*(1), 28. `http://doi.org/10.1525/collabra.91`