

# Appendix B

In dieser Datei wird eine vollständige Re-Implementation Reproduction von Studie 4 in Brotzeller & Gollwitzer (2025) nach den Schritten durchgeführt, die das *Conceptual Framework for Computational Reproductions* (Kohrt et al., 2024) vorgibt. Eine besser lesbare HTML-Version dieser Datei findet sich in den [Zusatzmaterialien](#) der vorliegenden Bachelorarbeit.

## (a) Defining the Scope

- **Works:** Studie 4 aus Brotzeller & Gollwitzer (2025)
- **Claims:**
  - H1 bestätigt: Größere Diskrepanzen zwischen Feedback und Selbstwahrnehmung ziehen größere Selbstkonzeptänderungen nach sich (S. 1739)
  - H2 widerlegt: Die wahrgenommenen Verbesserungsmöglichkeiten ziehen keine Asymmetrien in den Selbstkonzeptänderungen nach sich (S. 1739)
- **Results:** Tabelle 2 auf S. 1739. Diese Tabelle bildet alle Koeffizienten des Regressionsmodells mit den Prädiktoren `size of discrepancy` (SoD), `direction of discrepancy` (DoD) und `opportunity for improvement` (OFI) sowie all ihren Interaktionseffekten ab. Das Modell sagt die absoluten Selbstkonzeptänderungen vorher.
- **Obtained data:** Anonymisierte Primärdaten, die [hier](#) aus dem OSF abrufbar sind.
  - Die Primärdaten wurden schon so weit vorverarbeitet, dass Prädiktorvariablen des Regressionmodells bereits aus anderen Variablen im Datensatz berechnet wurden. Dieser Verarbeitungsschritt wird wiederholt.
  - Der Datensatz enthält noch alle Versuchspersonen, die Brotzeller & Gollwitzer (2025) aus der finalen Analyse ausschließen. Die Anwendung der präregistrierten Ausschlusskriterien kann und wird deshalb wiederholt werden.
- **Involved descriptions:** Die Beschreibungen aus den folgenden Quellen werden miteinbezogen und auf etwaige Inkohärenzen überprüft:

- Beschreibungen der Hypothesen, der Methoden, des Vorgehens und der Ergebnisse von Studie 4 im Artikel (S. 1738-1739)
- Beschreibungen der Schlussfolgerungen aus den Ergebnissen von Studie 4 in der Diskussion (S. 1740)
- Beschreibungen der Hypothesen, der geplanten Berechnung sowie der Ausschlusskriterien in der [Präregistrierung](#)
- Beschreibungen der Berechnung der für die Regressionsanalyse benötigten Variablen im Codebook
- Für die Formel zur Berechnung der Diskrepanz zwischen Feedback und Selbstwahrnehmung musste auf eine Beschreibung in den Materialien zu Studie 4 zurückgegriffen werden
- **Inherited source code:** Der Quellcode für die Berechnung wird re-implementiert. Es wird nur dann auf Teile des [Quellcodes der Autoren](#) zurückgegriffen, wenn eine offene Frage bei der Implementierung nicht eigenständig geklärt werden kann.

## (b) Checking the Preconditions

- Die anonymisierten [Primärdaten](#) konnten abgerufen und eingebunden werden:

```
original_data <- read.csv2("../data/CleanData_S4_final.csv")
```

- Die Autoren führen ihre Berechnung in R (R Core Team, 2018) durch. Für die Reproduktion wird die folgende R-Umgebung verwendet:

```
R version 4.5.2 (2025-10-31 ucrt)
x86_64-w64-mingw32/x64
Windows 11 x64 (build 26200)
```

- Die benötigten R-Pakete konnten installiert und eingebunden werden:

```
library(dplyr)
library(apaTables)
```

```
dplyr version 1.1.4
apaTables version 2.0.8
```

 **Fazit:** Die Reproduktion kann durchgeführt werden

Alle Voraussetzungen für die Durchführung der Reproduktion sind erfüllt.

## (c) Obtaining results

### Applying exclusion criteria

```
# We will exclude participants

# - if they fail our attention check item
# (i.e. if they don't select "strongly agree"
# as the responses despite being instructed to).
data_after_exclusion <- original_data %>% filter(attention_coping == 9)

# - if their responses to either of the discrepancy
# attention checks is incorrect. If this leads to
# more than 20% of participants being excluded, this
# exclusion criterion will be relaxed.
n_before <- nrow(data_after_exclusion)

data_after_exclusion <- data_after_exclusion %>%
  filter(attention_sc == "right" & attention_task == "right")

n_after <- nrow(data_after_exclusion)

if ((n_before - n_after) / n_before <= 0.2) {
  cat("No need to relax the exclusion criterion \
    \nregarding the discrepancy attention checks.")
}
```

No need to relax the exclusion criterion

regarding the discrepancy attention checks.

```
# - if they indicate that we should not use their data
# in their response to our "use-me" item.
data_after_exclusion <- data_after_exclusion %>% filter(use_me == "yes")
```

💡 Fazit: Ausschluss von Versuchspersonen erfolgreich reproduziert

Im Einklang mit der von Brotzeller & Gollwitzer (2025) berichteten Anzahl, wurden  $n = 89$  Personen von der Analyse ausgeschlossen.

## Data Pre-Processing

Der Datensatz wird auf die Spalten reduziert, die für die Berechnung der Variablen des Regressionsmodells benötigt sind:

```
data <- data_after_exclusion %>% select(  
  subject_nr,  
  condition_ofi,  
  starts_with("sc_t1_0"),  
  starts_with("sc_t2_0"),  
  task_emotion_score_percent  
)
```

Die Selbstwahrnehmungen zu den Zeitpunkten t1 und t2 errechnen sich aus dem Mittelwert der Antworten auf die vier Items, mit denen die Selbstwahrnehmung abgefragt wurde:

```
data <- data %>% mutate(  
  sc_t1 = rowMeans(across(c(sc_t1_01, sc_t1_02, sc_t1_03, sc_t1_04))),  
  sc_t2 = rowMeans(across(c(sc_t2_01, sc_t2_02, sc_t2_03, sc_t2_04))),  
  .before = task_emotion_score_percent # höhere Übersichtlichkeit  
)
```

Die absoluten Selbstkonzeptänderungen lassen sich aus dem Betrag der Differenz zwischen den Selbstwahrnehmungen zu beiden Zeitpunkten berechnen:

```
data <- data %>% mutate(sc_diff_absolute = abs(sc_t2 - sc_t1))
```

Die Größe der Diskrepanz (SoD) lässt sich aus dem Betrag der Differenz zwischen dem Feedback und der Selbstwahrnehmung zum Zeitpunkt t1 berechnen. Die Richtung der Diskrepanz (DoD) aus dem Vorzeichen dieser Differenz.

### ! Mistake 1: Inkohärenz bezüglich der Berechnung der Richtung der Diskrepanz

Im Codebook wird die Berechnung von DoD anders als im Artikel beschrieben. Die Beschreibung im Codebook ist falsch, weil hier die Diskrepanz zwischen Feedback und Selbstwahrnehmung aus der Differenz zwischen den Selbstwahrnehmungen zu beiden Zeitpunkten berechnet wird:

“negative”:  $(sc\_t2 - sc\_t1) < 0$

“neither positive nor negative”:  $(sc\_t2 - sc\_t1) = 0$

“positive”:  $(sc\_t2 - sc\_t1) > 0$

Weil das Feedback als Prozentwert vorliegt, muss die Differenz ebenfalls zu einem Prozentwert umgerechnet werden. Dazu müssen die Selbstwahrnehmungswerte von der ursprünglichen Skala von 1-9 auf die Skala 0-8 umgerechnet werden, die einen Nullpunkt hat. Das Feedback errechnet sich als prozentualer Gesamtwert aus den einzelnen Items des Reading the Mind in the Eyes Test (Bölte, 2005); diese Berechnung wird nicht noch einmal wiederholt, sondern von den Autoren übernommen.

```
# absolute discrepancy
data <- data %>% mutate(
  discrepancy = task_emotion_score_percent - ((sc_t1 - 1)/8)*100
)

# size of discrepancy
data <- data %>% mutate(sod = abs(discrepancy))

# direction of discrepancy
data <- data %>% mutate(dod = case_when(
  discrepancy > 0 ~ "positive",
  discrepancy < 0 ~ "negative",
  discrepancy == 0 ~ "neither positive nor negative"
))

# discrepancy ist keine Variable im originalen Datensatz
# und wird deshalb hier wieder entfernt.
data <- data %>% select(!discrepancy)
```

### ! Obstacle 1: Berechnung der Größe der Diskrepanz ist unterspezifiziert

Folgt man den Beschreibungen von Brotzeller & Gollwitzer (2025) bei der Berechnung von SoD und DoD, stimmen die Werte nicht mit jenen überein, die die Autoren in ihrem Datensatz berechnet haben.

Eine Inspektion des ursprünglichen Datensatzes hat ergeben, dass die Werte von SoD gerundet werden müssen, auch wenn das in keiner der Quellen Erwähnung findet. Auch die genaue Rundungsmethode wird nicht spezifiziert.

Erst nach einigen Fehlversuchen konnte die richtige Rundungsmethode identifiziert werden. Sie entspricht der klassischen Rundungsmethode, d.h. Werte unter .5 werden abgerundet, Werte darüber hingegen aufgerundet. Weil R keine Standardfunktion für diese Rundungsmethode implementiert, musste sie eigens definiert werden.

```
round_half_up <- function(x) {
  # floor(x) rundet immer auf den nächsten Integer-Wert ab.
  floor(x + 0.5)
```

```
}

data <- data %>% mutate(
  sod = abs(task_emotion_score_percent - round_half_up(((sc_t1 - 1)/8)*100))
)
```

💡 Fazit: Erstellen der Variablen des Regressionsmodells erfolgreich reproduziert

Nachdem Mistake 1 gelöst wurde und Obstacle 1 überwunden wurde, stimmen die Werte der neu berechneten Variablen mit den ursprünglichen Daten überein:

```
matches <- data %>%
  inner_join(data_after_exclusion)
```

In 459 von 459 Fällen stimmen die Daten überein.

## Regression analysis

Aus der Analyse werden die Daten aller Versuchspersonen ausgeschlossen, deren Selbstwahrnehmung mit dem Feedback übereinstimmt:

```
data <- data %>% filter(dod != "neither positive nor negative")
```

Von den 459 infragekommenden Versuchspersonen wurden zusätzliche n = 6 Personen von der Analyse ausgeschlossen.

Die Prädiktoren DoD und opportunity for improvement (OFI) werden effekt-kodiert:

```
data <- data %>% mutate(
  dod_num = case_match(
    dod,
    "negative" ~ -1,
    "positive" ~ 1
  )
)

data <- data %>% mutate(
  ec_condition_ofi = case_match(
    condition_ofi,
```

```

    "low opportunity for improvement" ~ -1,
    "high opportunity for improvement" ~ 1
  )
)

```

Um die Interpretierbarkeit des Modells zu erhöhen, wird SoD z-standardisiert:

```
data <- data %>% mutate(sod_standard = scale(sod))
```

Nun kann die Regressionanalyse durchgeführt werden. Dafür wird der `lm` Befehl verwendet, der in R standardmäßig implementiert ist:

```
model_fit <- lm(sc_diff_absolute ~ ec_condition_ofi*dod_num*sod_standard, data)
```

Die Ergebnisse der Regressionsanalyse können mit dem folgenden Befehl inspiziert werden:

```
summary(model_fit)
```

Call:

```
lm(formula = sc_diff_absolute ~ ec_condition_ofi * dod_num *
    sod_standard, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.9143	-0.6330	-0.1694	0.4057	3.8674

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.085588	0.046627	23.283	< 2e-16 ***
ec_condition_ofi	0.030227	0.046627	0.648	0.517
dod_num	-0.376120	0.046627	-8.067	6.77e-15 ***
sod_standard	0.384810	0.054055	7.119	4.39e-12 ***
ec_condition_ofi:dod_num	-0.035368	0.046627	-0.759	0.449
ec_condition_ofi:sod_standard	-0.001876	0.054055	-0.035	0.972
dod_num:sod_standard	-0.307136	0.054055	-5.682	2.41e-08 ***
ec_condition_ofi:dod_num:sod_standard	0.016376	0.054055	0.303	0.762

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8487 on 445 degrees of freedom  
Multiple R-squared: 0.1736, Adjusted R-squared: 0.1606  
F-statistic: 13.36 on 7 and 445 DF, p-value: 1.12e-15

---

Es fällt auf, dass der `lm`-Befehl für die einzelnen Koeffizienten nicht die Effektstärken  $sr^2$  (squared semi-partial correlations) berechnet. Da Brotzeller & Gollwitzer (2025) diese Effektstärken angeben, müssen sie nachträglich berechnet werden.

**i** Conflicting Choice 1: Mangelndes Wissen, um die Semipartialkorrelationen zu berechnen

Da ich weder weiß, wie man  $sr^2$  in R berechnet, noch, wie man diese Effektgröße händisch berechnet, muss ich an dieser Stelle die Berechnungsmethode aus dem Quellcode der Autoren übernehmen.

Der Befehl `apa.reg.table` aus dem Paket `apaTables` berechnet  $sr^2$  automatisch mit:

```
apa.reg.table(model_fit)
```

Regression results using `sc_diff_absolute` as the criterion

Predictor	b	b_95%_CI	sr2	sr2_95%_CI
(Intercept)	1.09**	[0.99, 1.18]		
ec_condition_ofi	0.03	[-0.06, 0.12]	.00	[-.00, .01]
dod_num	-0.38**	[-0.47, -0.28]	.12	[.07, .18]
sod_standard	0.38**	[0.28, 0.49]	.09	[.04, .14]
ec_condition_ofi:dod_num	-0.04	[-0.13, 0.06]	.00	[-.00, .01]
ec_condition_ofi:sod_standard	-0.00	[-0.11, 0.10]	.00	[-.00, .00]
dod_num:sod_standard	-0.31**	[-0.41, -0.20]	.06	[.02, .10]
ec_condition_ofi:dod_num:sod_standard	0.02	[-0.09, 0.12]	.00	[-.00, .00]

Fit



```
R2 = .174**  
95% CI[.10,.22]
```

Note. A significant b-weight indicates the semi-partial correlation is also significant.  
b represents unstandardized regression weights.  
sr2 represents the semi-partial correlation squared.  
Square brackets are used to enclose the lower and upper limits of a confidence interval.  
\* indicates  $p < .05$ . \*\* indicates  $p < .01$ .

---

#### (d) Evaluating the Consistency of Results

💡 Fazit: Die berichteten Ergebnisse konnten reproduziert werden

Ein manueller Vergleich der Neuberechneten mit den berichteten Ergebnissen zeigt, dass die Reproduktion **konsistente** Ergebnisse hervorgebracht hat. Rundet man die Neuberechneten Werte auf zwei Nachkommastellen, sind sie sogar **numerisch identisch** zu den in Tabelle 2 (S. 1739) enthaltenen Werten sind.