10.20378/irb-108746



University of Bamberg

# New insights and approaches for dealing with response biases in psychological assessment

Dr. habil. Jessica Röhner

https://www.uni-bamberg.de/perspsych/team/jessica-roehner/

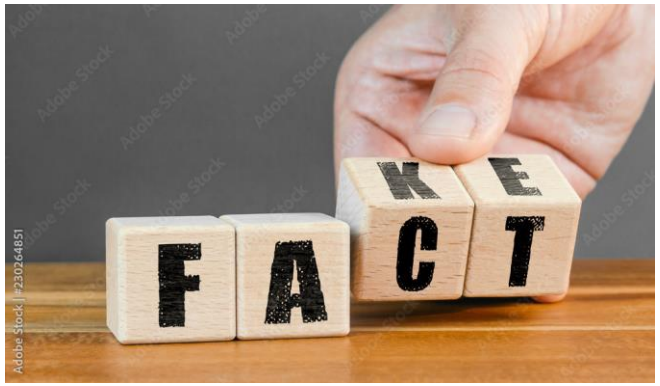
jessica.roehner@uni-bamberg.de

# Why is response distortion a relevant topic?

- Task of psychological assessment: Formulation of ***empirically based and, as far as possible, correct estimates*** of the status of psychological constructs and their change in individuals (e.g. Beauducel & Leue, 2014; Röhner & Schütz, 2022)
- ***Description***, ***classification***, ***explanation***, ***prognosis, and evaluation*** of states and processes (e.g., Beauducel & Leue, 2014)
- Used to obtain ***relevant information***
- Used to ***answer questions*** based on this information
- Example: Prognosis regarding the likelihood of a future criminal offense in an imprisoned offender

=> In most cases, it is of great importance that the psychological assessment is as accurate as possible because human destinies often depend on the recommendations (e.g., recommendations regarding school transfer, professional suitability, custody of a child, or early release from prison)

# Why is response distortion a relevant topic?



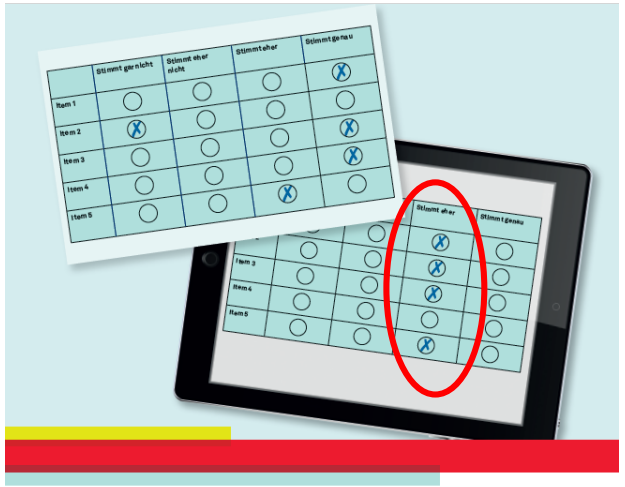

- The aim of using psychological assessment procedures is to asses the ***characteristics of the examined persons as correctly as possible*** (Röhner & Schütz, 2025)
- Requirement: only the ***characteristic of interest is measured*** and nothing else (concept of response validity)

⇒ Response validity = extent to which the data collected (e.g., in the form of responses in self-report scales) from the investigated individuals reflect their actual attributes (e.g., attitudes) (Edwards, 2019)

⇒ Component of construct validity (e.g. Huang et al., 2015)

# Why is response distortion a relevant topic?

- Over the course of conducting psychological assessments, practitioners and researchers will most likely encounter the phenomenon of response bias
- In the literature, many different terms are associated with response distortion:
  - Lies
  - Social desirability
  - Deceptive self-presentation
  - Faking
  - Carelessness
  - Self-deception
  - Impression management
  - Simulation
  - Dissimulation



Jessica Röhner
Astrid Schütz

Phänomene der Antwortverzerrung in der Psychologischen Diagnostik

# Agenda


1. ***Typical examples of response biases and their facets***
    1. Carlessness
    2. Socially desirable responding (SDR)
    3. Faking
2. ***Frequencies of response biases***
3. ***Effects*** (e.g., on means, reliability, and validity) ***and consequences*** (e.g., erroneous decisions in assessment) ***of response biases***
4. ***Strategies to prevent and detect/correct response biases***
5. ***Current research***
6. ***Discussion***

# 1. Typical examples of response biases and their facets

# 1. Typical examples of response biases and their facets

- ***Response biases*** = systematic biases in the responses of participants (Paulhus, 1991; Moosbrugger & Brandt, 2020)
- If response distortions occur, the answers of tested individuals in the psychological assessment no longer reflect the characteristic to be measured only, but also contain other sources of influence (e.g., whether the person generally tends to agree; i.e., *acquiescence*)

- Response biases:
  - ***Response styles*** (biases that are consistent across time and measures; e.g., *extreme responding*)
  - ***Response sets*** (biases that vary with time and across measures; caused by temporary distraction or change in motivation; e.g., *faking*)

# 1. Typical examples of response biases and their facets

## Carelessness

- Key feature: Largely ignoring the test instructions and/or item content
- Cause: Lack of motivation, not lack of ability to respond carefully (e.g. Goldammer et al., 2020; Schroeders et al., 2022)
- Careless respondents, for example, skip items (e.g., only 5 of 10 items are answered) or answer in patterns (e.g., "Solution A," "Solution C," "Solution A," "Solution C,"....)
- Typical facets are:
    1. Omitting responses
    2. Consecutive identical responses
    3. „Random“ responding
    4. Patterned responding

# 1. Typical examples of response biases and their facets

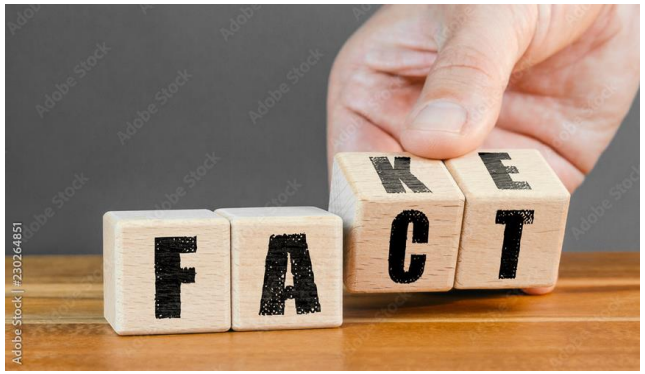
## Socially desirable responding (SDR)

- Key feature: Understood as a tendency towards positive self-description (Paulhus, 2002)
- <u>Typical facets are:</u>
    1. Self-deceptive enhancement (having a positive impression of oneself)
    2. Impression management (trying to impress others)
    3. Agentic impression management (self-presentation as a hero)
    4. Communal impression management (self-presentation as a saint)

# 1. Typical examples of response biases and their facets

## Faking

- Key feature: Describes the intentional distortion of responses in order to achieve a specific goal (e.g., a job offer; Ziegler et al., 2011; see also Röhner & Schütz, 2019)
- In addition to faking motivation, the ability to fake, and the opportunity to fake are necessary preconditions
- Typical facets are:
    1. Faking good versus faking bad
    2. Simulation versus dissimulation
    3. Slight faking versus extreme faking
    4. Blatant extreme responding versus intermittent faking



# 2. Frequencies of response biases

# 2. Frequencies of response biases

## Carelessness

- 1% of respondents are careless responders (Gough & Bradley, 1996)
- 78% of respondents are careless responders (Mancosu et al. 2019)

## SDR

- At least 30 % of respondents are socially desirable responders (Ambwani et al., 2013; Holden & Book, 2011)

## Faking

- 10 % of respondents are fakers (Pella et al., 2012)
- 62 % of respondents are fakers (Donovan et al., 2003)

=> Base rate varies concerning the investigated samples (e.g., click workers vs. students), settings (e.g., high-stakes vs. low-stakes setting), and detection methods (e.g., index for straightlining vs. index for omitted responses)

# 3. Effects and consequences of response biases

# 3. Effects and consequences of response biases

## Effects

- Response biases impact:
  - Means (e.g., carelessness: Goldammer et al., 2020 and Rios et al., 2017; SDR: Bäckström & Björklund, 2013; faking; Birkeland et al., 2006 and Salgado, 2016)
  - Correlations (e.g., carelessness: Huang et al., 2015 and Jaso et al., 2021; SDR: Bäckström et al., 2009; Connelly & Chang, 2016; faking; Bressan et al., 2018 and Salgado, 2016)
  - Reliability estimates (e.g., carelessness: DeSimone et al., 2018 and Maniaci & Rogge, 2014; SDR: Bäckström et al., 2009 and Wood et al., 2022; faking; Röhner et al., 2022 and Salgado, 2016)
  - Validity estimates (e.g., carelessness: Arthur et al., 2021 and Maniaci & Rogge, 2014; SDR: Kallio Strand et al., 2021 and Paunonen & LeBel, 2012; faking; Dunlop et al., 2022 and Ziegler & Bühner, 2009)
  - Factor structure (e.g., carelessness: Johnson, 2005 and Woods, 2006; SDR: Bäckström & Björklund, 2020 and Jackson et al., 1996; faking; Ellingson et al., 1999; Van Iddekinge et al., 2005)
  - Rank ordering (e.g., faking; Ellingson et al., 1999; Griffith et al., 2007 => In a sample of 60 people, some interviewees dropped 48 positions under honest conditions compared to the application situation)

## Consequences

- Response biases impact:
  - Theory development (e.g., carelessness: Arthur et al., 2021; SDR: Musch, 1999; faking; MacCann et al., 2011)
  - Decision Making (e.g., in forensic settings or in personnel selection; e.g., SDR: Hildebrand et al., 2018; faking; Burns et al., 2017)

# 4. Strategies to prevent and detect/correct response biases

# 4. Strategies to prevent and detect/correct response biases

**Proactive versus reactive strategies**

- Proactive strategies: used to prevent the biases
- Reactive strategies: used to detect (and correct) the biases

Because research concerning valid indicators for response biases is lacking for most measures, it is recommendable to favor proactive strategies over reactive ones.

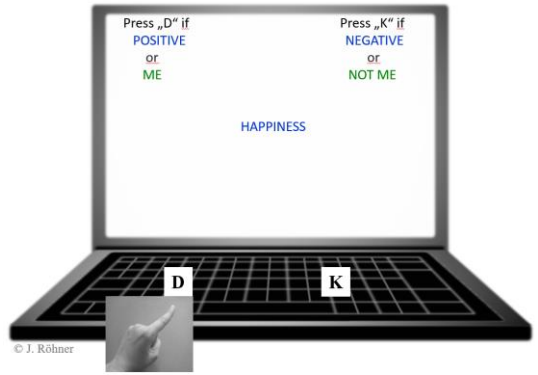
# 4. Strategies to prevent and detect/correct response biases

Typical proactive and reactive strategies taking faking as an example

## Proactive strategies

1. ***Implementation of warnings***
   - Several types of warnings (e.g., warnings that faking can be identified, warnings that faking is morally unacceptable; Dilchert & Ones, 2011)
   - Indications of reduced faking after warnings on possible identification (Law et al., 2016)
   - BUT:
     - Respondents just apply more sophisticated faking strategies (Youngjohn et al., 1999)
     - Generalized warnings to all respondents are ineffective and may even have negative side effects (Feeney et al., 2023; e.g., cause anxiety in honest responders: Li et al., 2022)
     - Ethic concerns because valid faking detection is still not possible for all measures (so one has to lie to participants)

# 4. Strategies to prevent and detect/correct response biases




Typical proactive and reactive strategies taking faking as an example

## Proactive strategies

***2. Using measures that are “immune” against faking***

- Several measures have been suggested to be immune against faking
- Most prominently, Implicit Association Tests (IATs; Greenwald et al., 1998) have been suggested to be fake-resistant measures (but see Röhner & Iliescu, 2023)
  - Because they do not rely on self-reports of respondents but on reaction times (and errors)
  - Because they were considered to assess implicit measures that were assumed to be uncontrollable by respondents (but see also Corneille & Hütter, 2020; Corneille & Lush, 2023; De Houwer, 2006)

=> However, not only have IATs been demonstrated to be fakeable (e.g., Fiedler & Bluemke, 2005; Röhner et al., 2011) but also it has been shown that respondents use a variety of faking strategies to fake on IATs (Röhner et al., 2023)

# 4. Strategies to prevent and detect/correct response biases

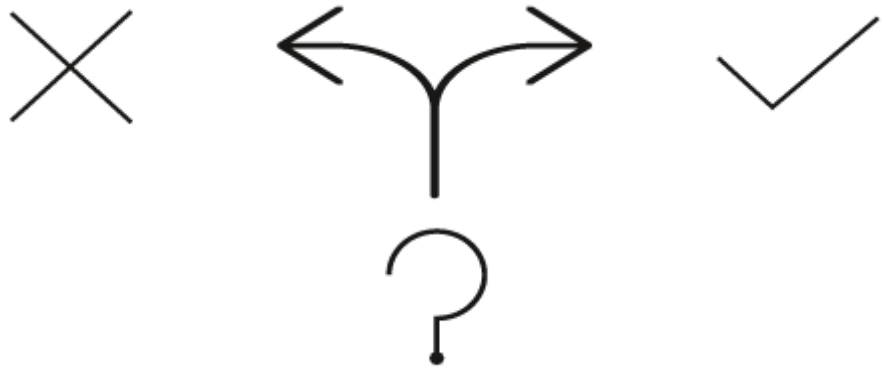
Typical proactive and reactive strategies taking faking as an example

**Proactive strategies**

***3. Using forced-choice formats***

***4. Reducing response latencies***

**…**



# 4. Strategies to prevent and detect/correct response biases

Typical proactive and reactive strategies taking faking as an example

**Reactive strategies:**

1. ***Using socially desirability scales to detect faking***
2. ***Applying faking indices to detect faking***
3. ***Analyzing participants' response latencies to detect faking***
4. ***Using AI to detect faking***
5. **…**

# 5. Current research

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

Cooperation with Astrid Schuetz and Ronald R. Holden

## Theoretical Background

- Indices [slowing, speeding, increasing or reducing errors in congruent or incongruent blocks; Combined Task Slowing (CTS); Ratio 150-10000] allegedly detect faking in IATs (Agosta et al., 2011; Cvencek et al., 2010; Röhner et al., 2013)
- However, studies are largely underpowered and inconclusive (Röhner et al., 2023)

# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRAVERSION | Talkative<br>Shy | INTROVERSION |
| **Practice** | ME | I<br>They | NOT ME |
| **Test (Compatible condition)** | EXTRAVERSION<br>or<br>ME | Talkative<br>I<br>Shy<br>They | INTROVERSION<br>or<br>NOT ME |
| **Practice** | NOT ME | They<br>I | ME |
| **Test (Incompatible condition)** | EXTRAVERSION<br>or<br>NOT ME | Talkative<br>They<br>Shy<br>I | INTROVERSION<br>or<br>ME |

# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRAVERSION | Talkative<br>Shy | INTROVERSION |
| **Practice** | ME | I<br>They | NOT ME |
| **Test (Compatible condition)** | EXTRAVERSION<br>or<br>ME | Talkative<br>I<br>Shy<br>They | INTROVERSION<br>or<br>NOT ME |
| **Practice** | NOT ME | They<br>I | ME |
| **Test (Inompatible condition)** | EXTRAVERSION<br>or<br>NOT ME | Talkative<br>They<br>Shy<br>I | INTROVERSION<br>or<br>ME |

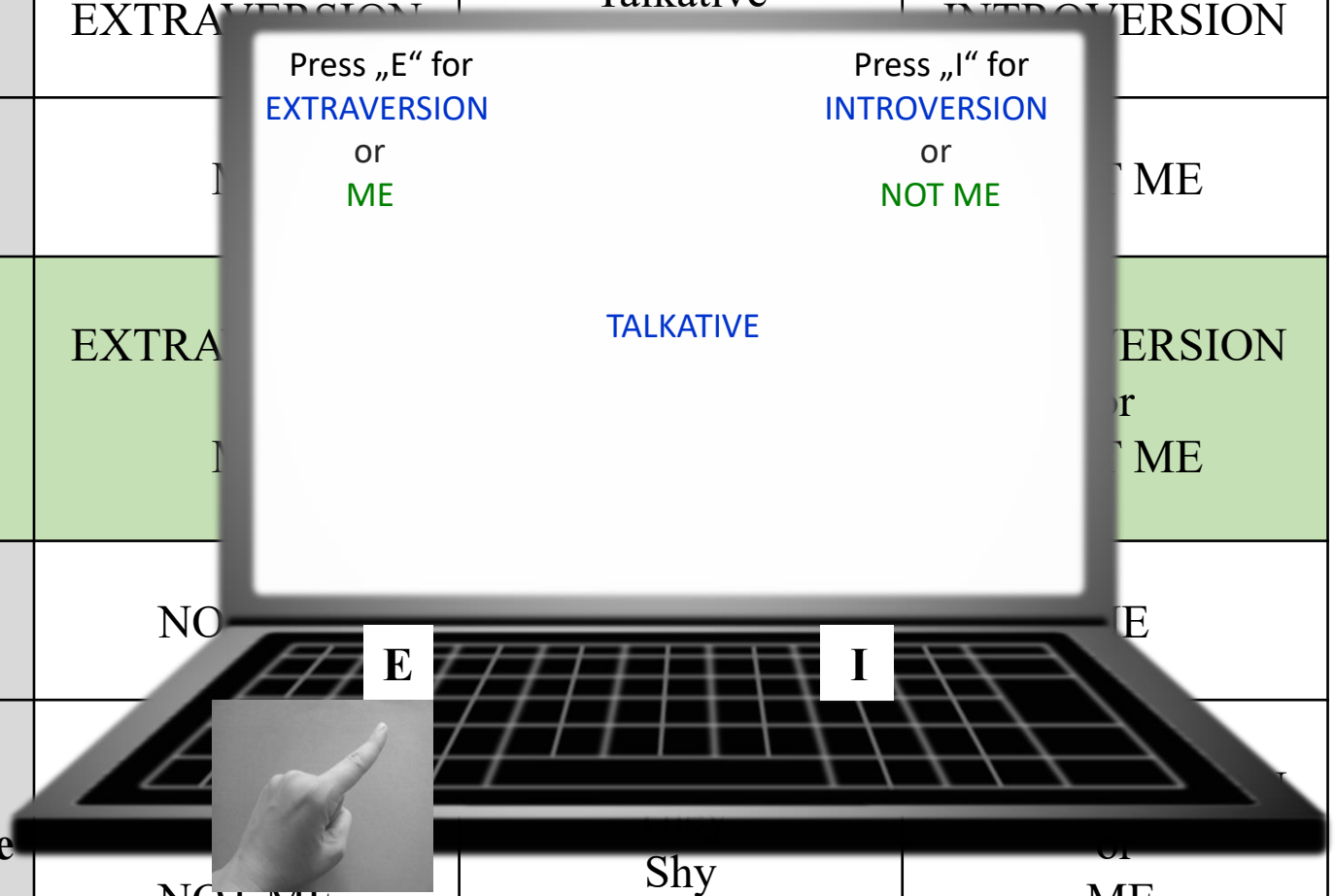
Not relevant for the computation of IAT effects

# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRAVERSION | Talkative<br>Shy | INTROVERSION |
| **Practice** | ME | I<br>They | NOT ME |
| **Test (Compatible condition)** | EXTRAVERSION<br>or<br>ME | Talkative<br>I<br>Shy<br>They | INTROVERSION<br>or<br>NOT ME |
| **Practice** | NOT ME | They<br>I | ME |
| **Test (Incompatible condition)** | EXTRAVERSION<br>or<br>NOT ME | Talkative<br>They<br>Shy<br>I | INTROVERSION<br>or<br>ME |

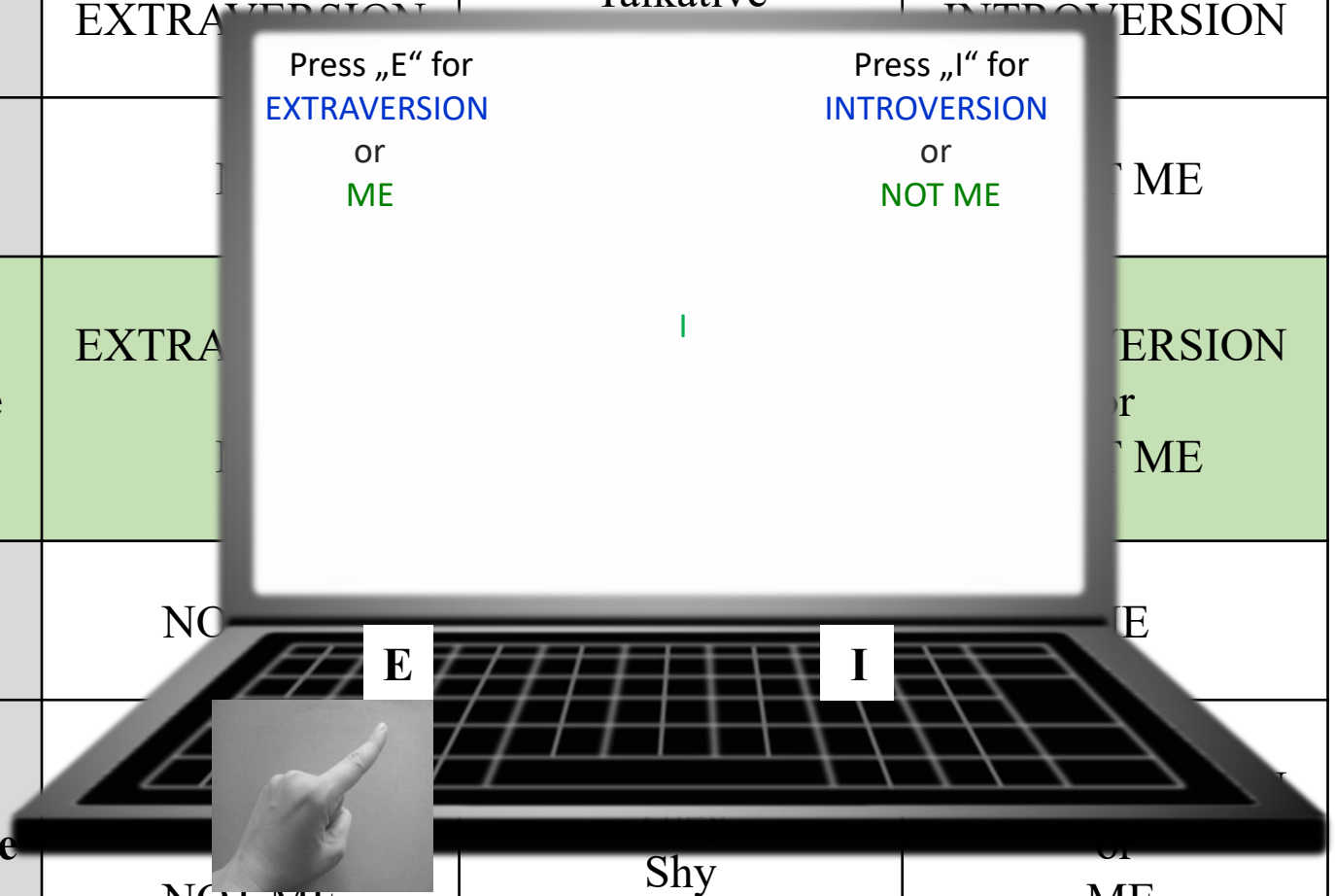
# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |



# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | I | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |

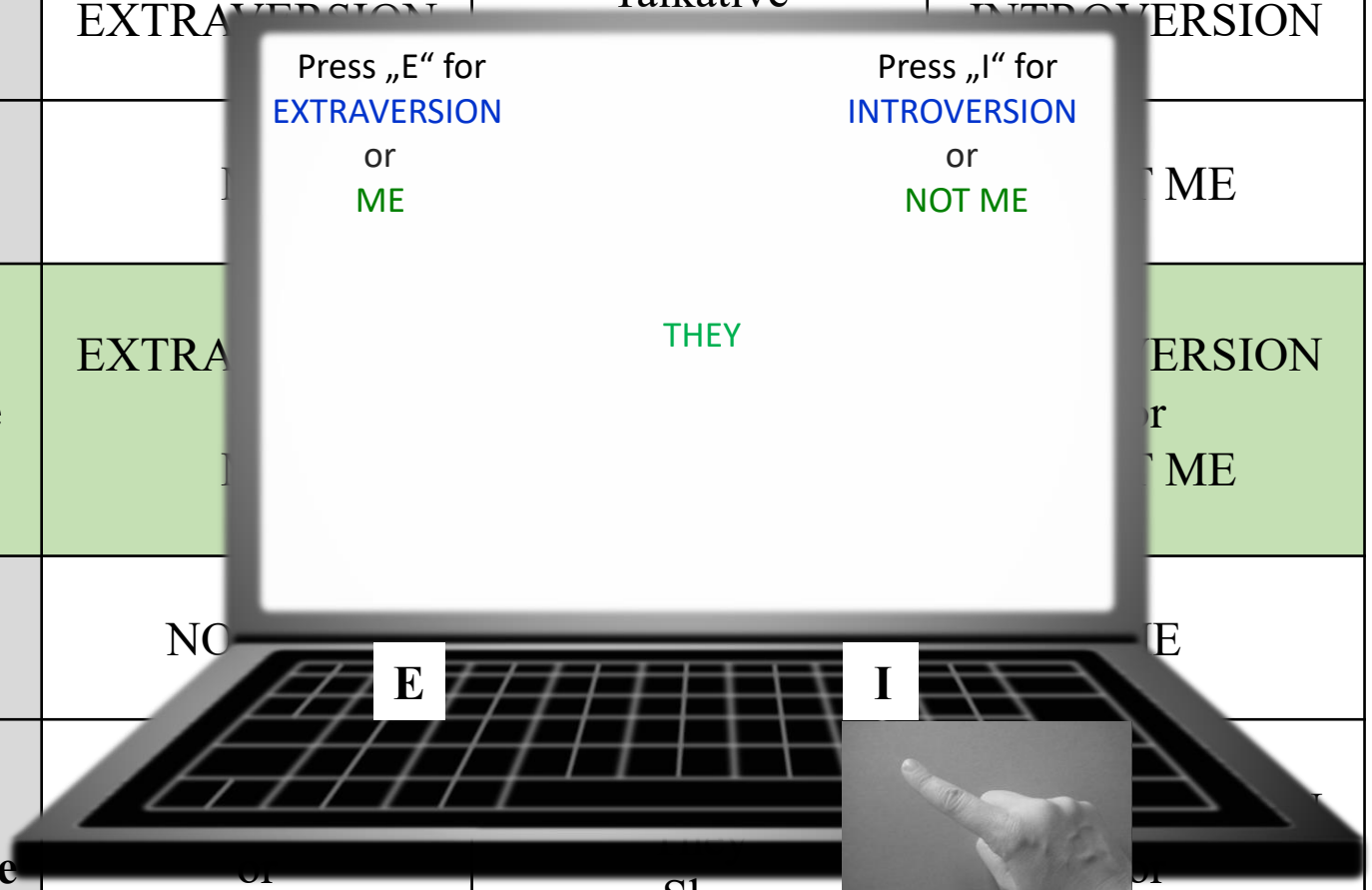


# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |

# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |

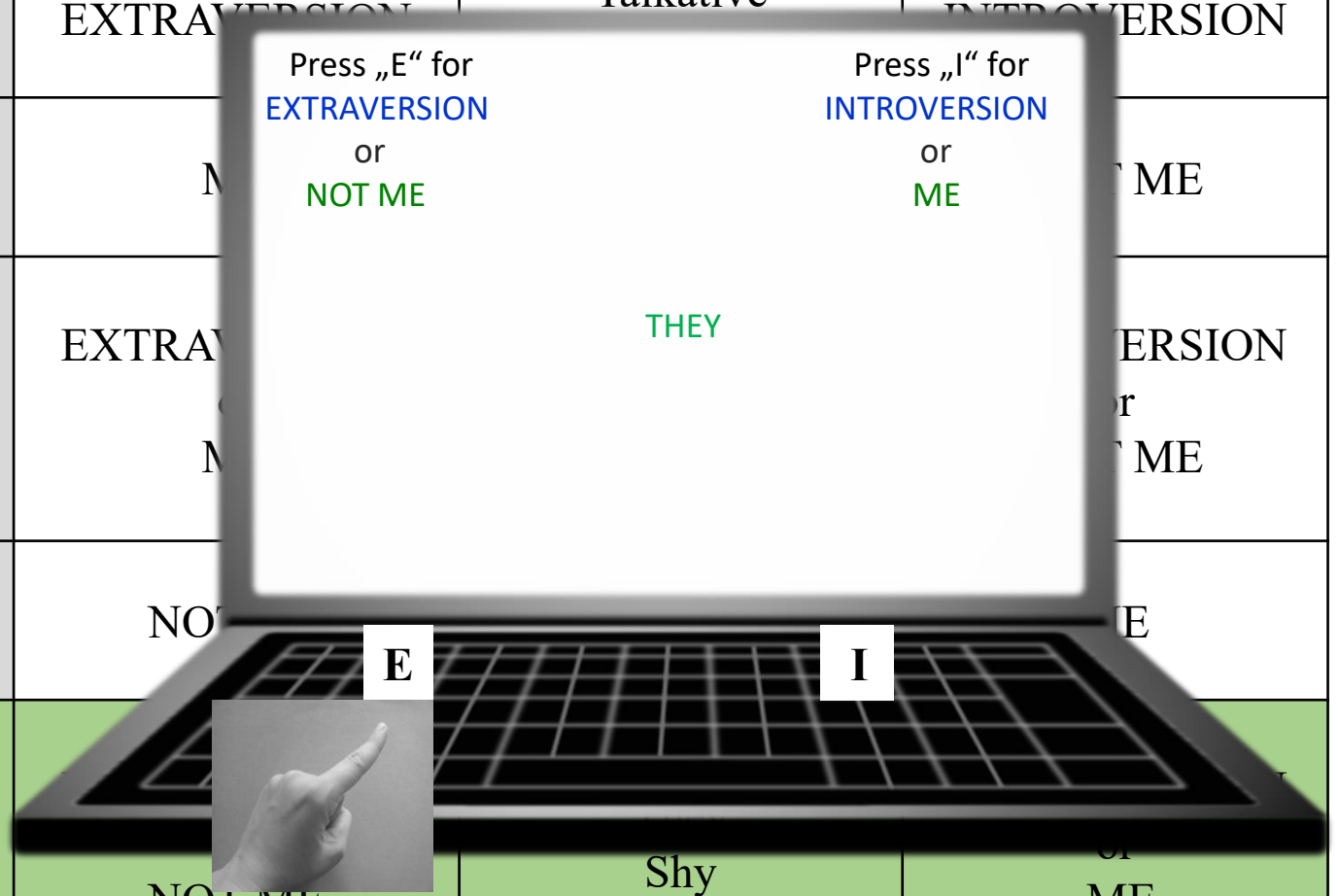


# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |

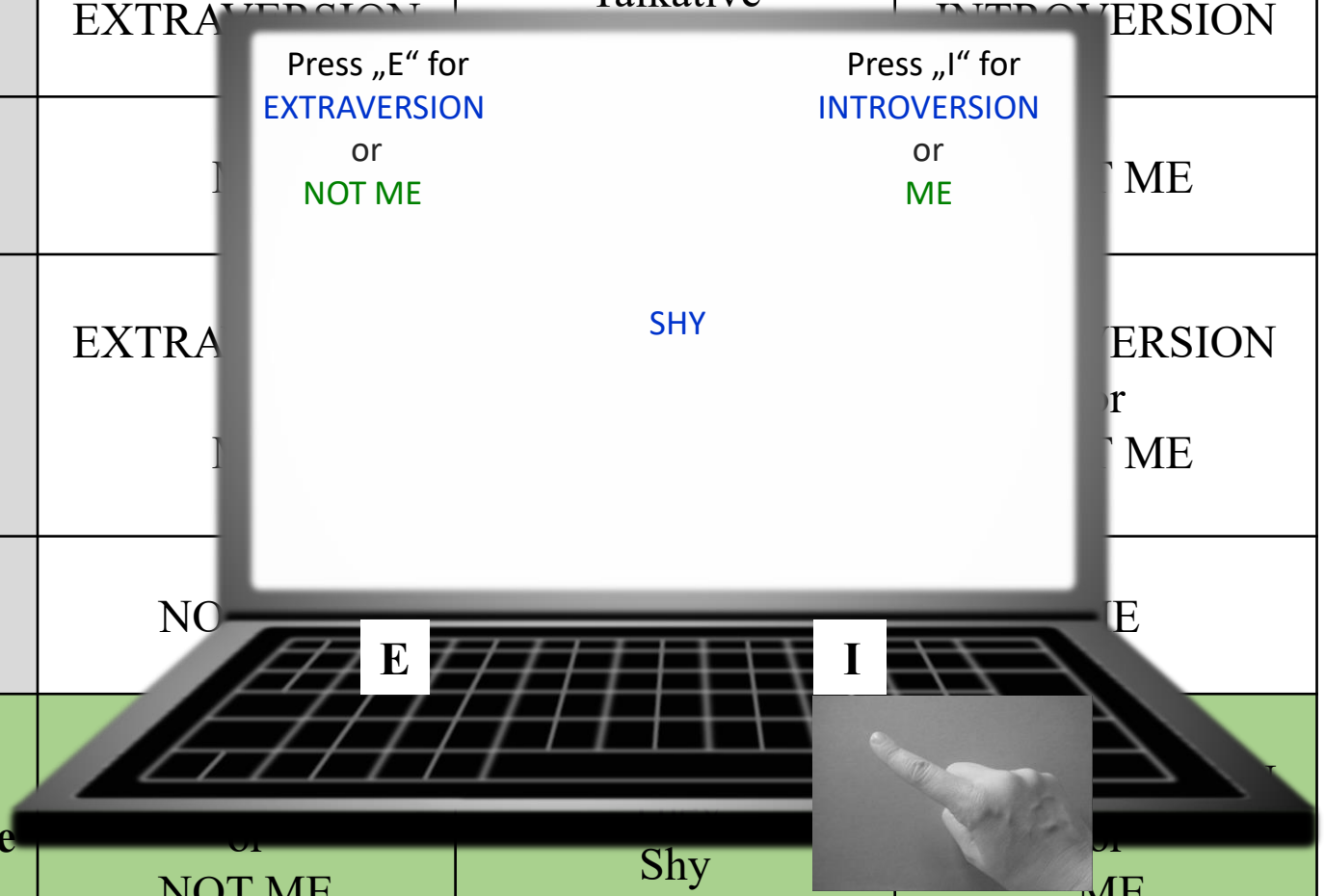
# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | M | | ME |
| **Test (Compatible condition)** | EXTRA<br>M | | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Inompatible condition)** | NOT ME | Shy<br>I | ME |



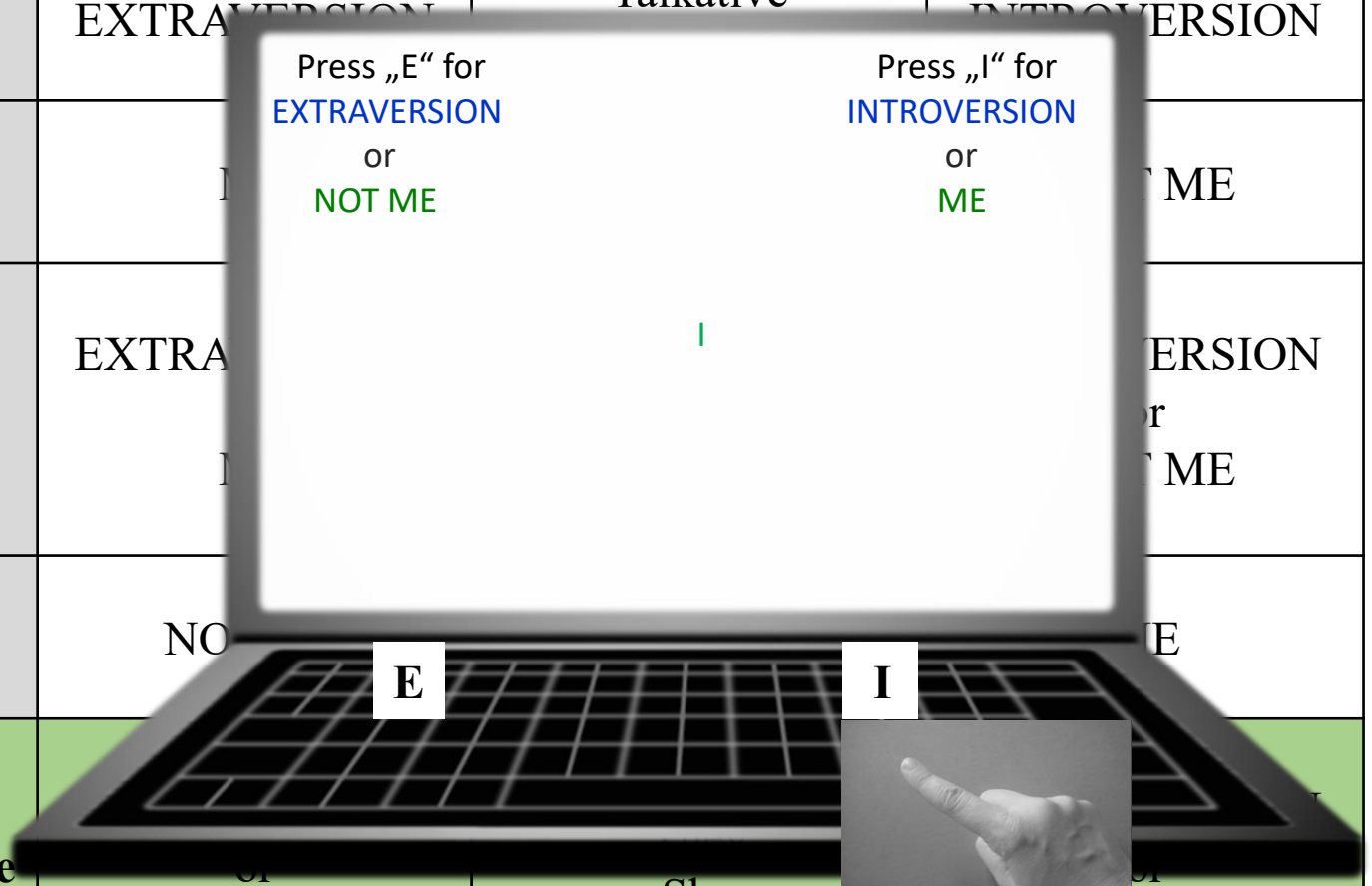
# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |



# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRA | Talkative | ERSION |
| **Practice** | | | ME |
| **Test (Compatible condition)** | EXTRA | I | ERSION<br>r<br>ME |
| **Practice** | NO | | E |
| **Test (Incompatible condition)** | NOT ME | Shy<br>I | ME |



# The Procedure of IATs (example: extraversion IAT)

| IAT Block | Left Key (*e*) | Example Stimuli | Right Key (*k*) |
|---|---|---|---|
| **Practice** | EXTRAVERSION | Talkative<br>Shy | INTRODUCTION |
| **Practice** | ME | I<br>They | NOT ME |
| **Test (Compatible condition)** | EXTRAVERSION<br>or<br>ME | Talkative<br>I<br>Shy | INTROVERSION<br>or<br>NOT ME |
| **Test (Incompatible condition)** | EXTRAVERSION<br>or<br>NOT ME | Talkative<br>They<br>Shy<br>I | INTROVERSION<br>or<br>ME |

$$\textbf{IAT effekt} = \frac{\text{RT incompatible IAT block} - \text{RT compatible IAT block}}{\textit{SD} \text{ in RT}}$$

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Faking Strategies and Faking Indices

| IAT block | Faking goal | |
|---|---|---|
| | **Low scores** | **High scores** |
| Faking indices that are based on conceptually derived faking strategies (Röhner et al., 2013) | | |
| Congruent | *Slowing down on the congruent block* (i.e., Slow_Co; difference in reaction time between the congruent block under faking and the congruent block at baseline)<br>*Increasing errors on the congruent block* (i.e., IncErr_Co; difference in errors between the congruent block under faking and the congruent block at baseline) | *Acceleration on the congruent block* (i.e., Accel_Co; difference in reaction time between the congruent block at baseline and the congruent block under faking)<br>*Reducing errors on the congruent block* (i.e., RedErr_Co; difference in errors between the congruent block at baseline and the congruent block under faking) |
| Incongruent | *Acceleration on the incongruent block* (i.e., Accel_In; difference in reaction time between the incongruent block at baseline and the incongruent block under faking)<br>*Reducing errors on the incongruent block* (i.e., RedErr_In; difference in errors between the incongruent block at baseline and the incongruent block under faking) | *Slowing down on the incongruent block* (i.e., Slow_In; difference in reaction time between the incongruent block under faking and the incongruent block at baseline)<br>*Increasing errors on the incongruent block* (i.e., IncErr_In; difference between the incongruent block under faking and the incongruent block at baseline) |
| Faking indices that are based on slowing behavior | | |
| Faster and slower | *CTS* (i.e., difference between the slower IAT block under faking and the faster IAT block under non-faking; Cvencek et al., 2010) | |
| Single and faster | *Ratio 150–10000* (i.e., ratio between the faster IAT block and the single IAT blocks under faking; Agosta et al., 2011) | |

Table 1

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Hypotheses

1) *Faking detection in faking low and in faking high conditions would differ with respect to the faking indices*.

2) *Not all strategies that are implemented by fakers are successful*.

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Participants and Data Sets

- Three data sets with extraversion IATs
- Final sample: 750 participants (258 faking low, 245 control, 247 faking high; 576 women, 173 men, 1 no response; 744 students); average age of 22.05 years (*SD* = 4.07)

## Procedure

- Participants took part in exchange for feedback and/or partial credit
- In all studies, participants completed the extraversion IAT twice
- On the first occasion (i.e., baseline), participants completed the IAT under standard instructions
- On the second occasion, participants were randomly assigned to one of three conditions (i.e., control, faking high scores, or faking low scores)
- Participants in the control condition again responded under standard instructions on the IAT
- Fakers were asked to fake either high scores or low scores on the IAT according to a personnel selection scenario

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Extraversion-IAT (Back et al., 2009)

- Self-relevant words (e.g., I, self) and non-self-relevant words (e.g., they, yours); extraversion-related words (e.g., talkative, active) and introversion-related words (e.g., shy, passive)

## Faking Strategies and Faking Indices

- Computed as described in Table 1

## Analytical Strategy

- ANOVA with repeated measures on the extraversion IAT *D* scores as a manipulation check to investigate whether participants in the faking groups were able to fake the IAT (was successful)
- ROC curve analyses to evaluate how well each of the strategies predicted whether participants belonged to the control group or a faking group
- Correlation analyses to evaluate how strategies were related to faking success
- Fisher's z test (Fisher, 1950) to compare the correlations in the faking groups to those in the control group

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Main Findings Concerning the Implementation and the Success of Faking Strategies

| Faking strategies and indices | Implementation | | Correlation with | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Faking success | | | Effects of repeated measurement | | |
| | | | When computed as *D* change | | | | | |
| | *AUC* | *SE* | *r* | *p* | *n* | *r* | *p* | *n* |
| Slowing down on the congruent block | **.84** | .02 | **.51** | < .001 | 253 | .36 | < .001 | 241 |
| Acceleration on the incongruent block | .37 | .03 | .41 | < .001 | 256 | .47 | < .001 | 243 |
| Increasing errors on the congruent block | **.84** | .02 | .03 | .587 | 258 | .20 | .001 | 243 |
| Reducing errors on the incongruent block | .29 | .02 | **.23** | < .001 | 258 | .05 | .422 | 244 |
| CTS | **.82** | .02 | **.17** | .006 | 253 | -.11 | .089 | 243 |
| Ratio 150–10000 | .45 | .03 | -.12 | .055 | 257 | .15 | .022 | 245 |

Faking low scores

*AUC*s in bold indicate that the strategy- or index-classified participants as belonging to the control or faking low group at levels above chance (> .50). Faking success = changes in IAT effects according to faking instructions. Effects of repeated measurement = changes in IAT effects in the control group (i.e., not due to faking instructions). Correlations printed in bold indicate that the significant positive correlation between the relevant faking strategy or faking index and faking success in the faking low group was significantly higher than the correlation between the respective behavior and effects of repeated measurement in the control group according to Fisher's *z* tests at $p < .05$.

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Main Findings Concerning the Implementation and the Success of Faking Strategies

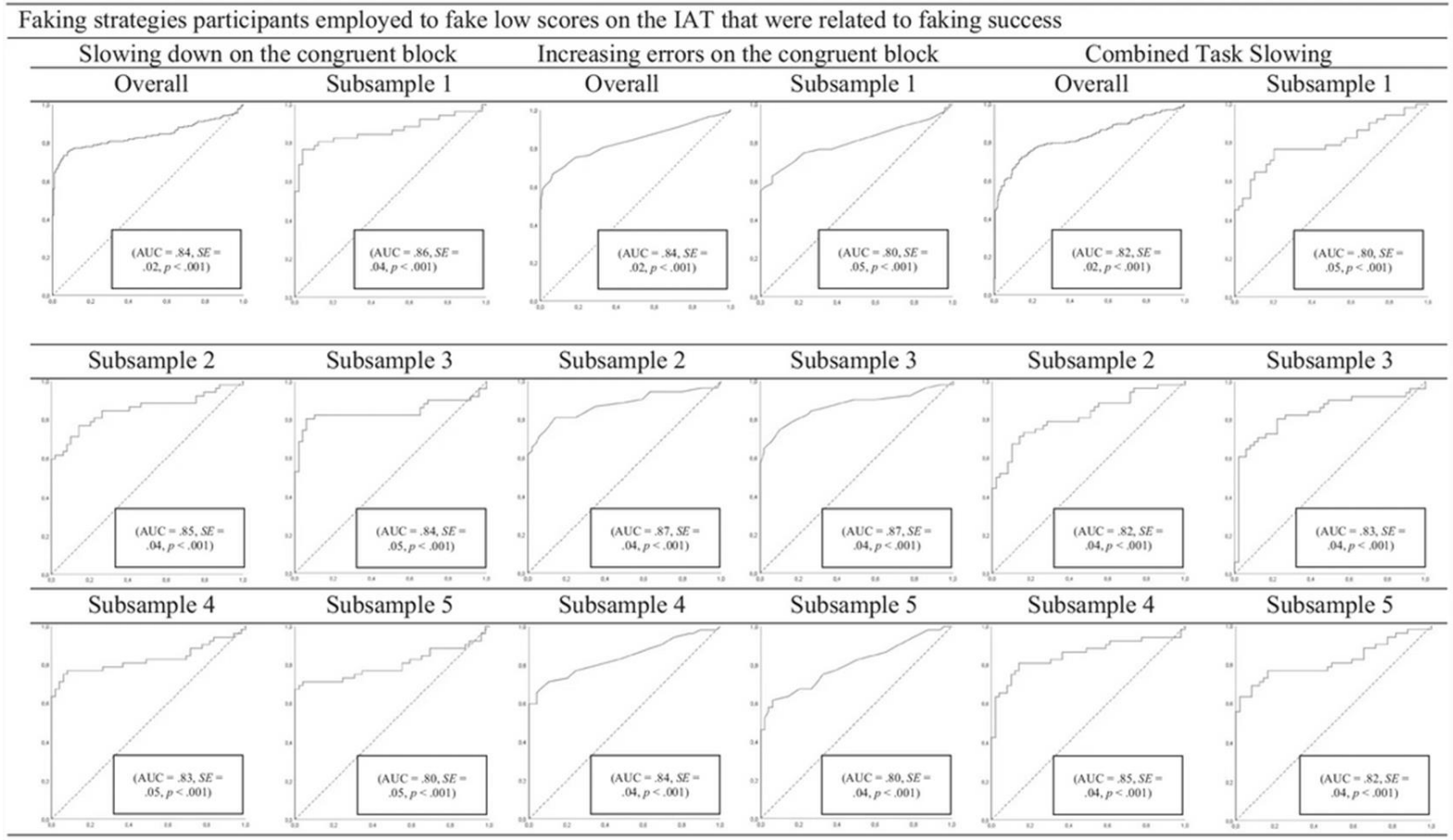
| Faking strategies and indices | Implementation | | Correlation with | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Faking success | | | Effects of repeated measurement | | |
| | | | When computed as *D* change | | | | | |
| | *AUC* | *SE* | *r* | *p* | *n* | *r* | *p* | *n* |
| Slowing down on the incongruent block | **.70** | .02 | **.64** | < .001 | 246 | .47 | < .001 | 243 |
| Acceleration on the congruent block | .51 | .03 | .41 | < .001 | 244 | .36 | < .001 | 241 |
| Increasing errors on the incongruent block | **.70** | .02 | **.37** | < .001 | 246 | .05 | .422 | 244 |
| Reducing errors on the congruent block | .49 | .03 | .16 | .016 | 236 | .20 | .001 | 243 |
| CTS | **.70** | .02 | **.36** | < .001 | 246 | .11 | .089 | 243 |
| Ratio 150–10000 | .43 | .03 | −.13 | .050 | 245 | −.15 | .022 | 245 |

Faking high scores

*AUC*s in bold indicate that the strategy or index classified participants as belonging to the control or faking high group at levels above chance (> .50). Faking success = changes in IAT effects according to faking instructions. Effects of repeated measurement = changes in IAT effects in the control group (i.e., not due to faking instructions). Correlations printed in bold indicate that the significant positive correlation between the relevant faking strategy or faking index and faking success in the faking high group was significantly higher than the correlation between the respective behavior and effects of repeated measurement in the control group according to Fisher's *z* tests at $p < .05$.

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

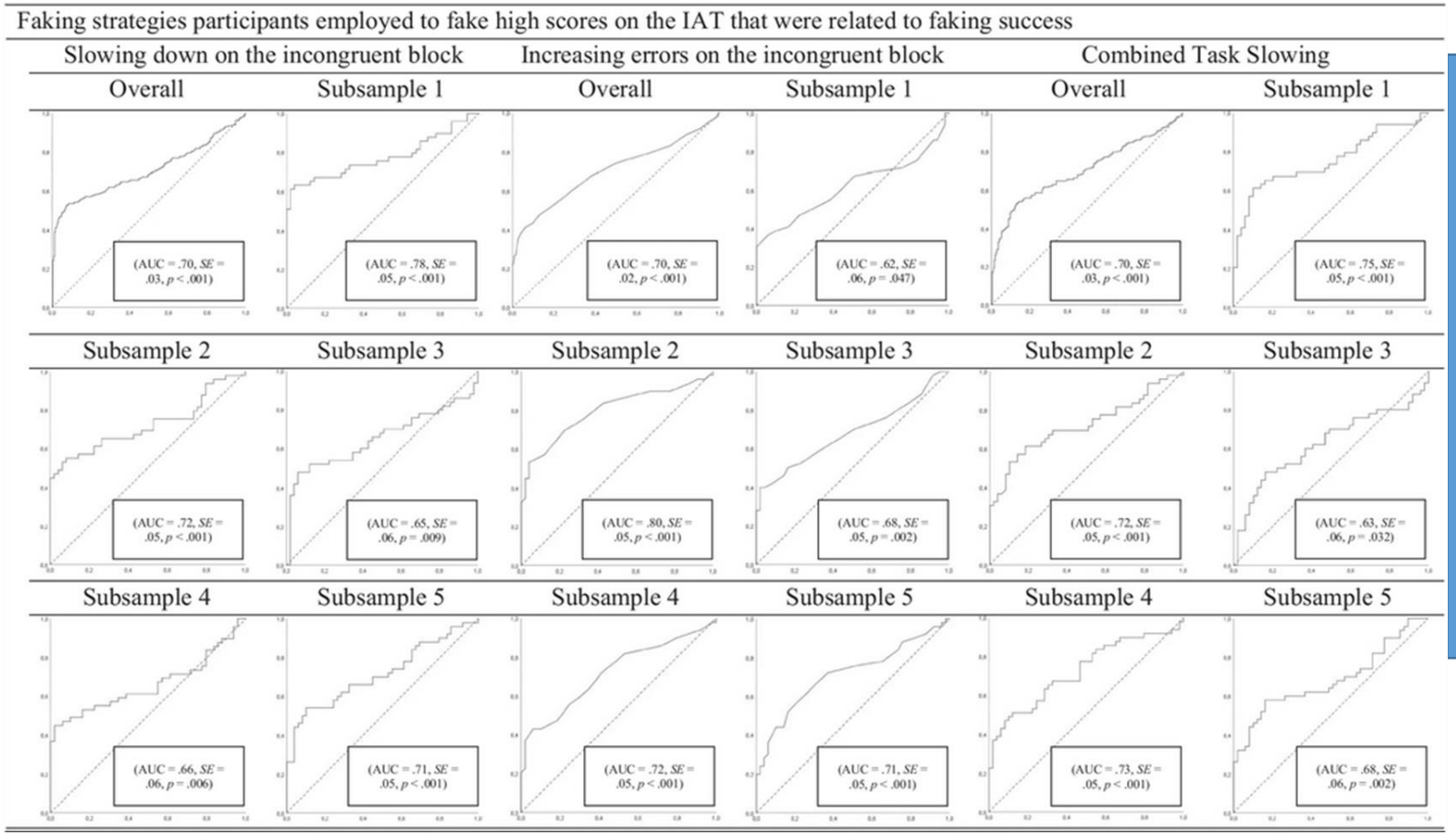
## Stability of Faking Indices




Faking low scores

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Stability of Faking Indices




Faking high scores

# IAT Faking Indices Revisited: Aspects of Replicability and Differential Validity

## Results in a Nutshell

1) Fakers use different faking strategies when faking low scores than when faking high scores
2) Not all faking indices are successful at levels above chance
3) Results are stable with respect to subsamples
4) Not all behaviors that revealed faking were successful in changing IAT effects as desired
5) Not all behaviors that were successful in changing IAT effects as desired revealed faking

## Conclusion and Implications

- Apparently, fakers use goal-dependent strategies which are not necessarily successful
- To detect faking, we recommend combining indices and considering the context

# Other Research Projects

- **Using AI to detect faking**: Lying on the dissection table: Anatomizing faked responses (Röhner et al., 2022)
- **Applying diffusion models to (faked) IAT data**: A diffusion model approach for understanding the impact of 17 interventions on the race Implicit Association Test (Röhner & Lai, 2021)
- **Applying faking models to new assessment conditions (few items, no warnings)**: Challenging response latencies in faking detection: The case of few items and no warnings (Röhner & Holden, 2022)
- **Trying to find out what people do to fake on self-reports**: Faking in self-report personality scales: A qualitative analysis and taxonomy of the behaviors that constitute faking strategies (Röhner, Schütz, & Ziegler, 2025)
- **Trying to distinguish faking effects from frame of reference effects**: A registered report to disentangle the effects of frame of reference and faking in the personnel-selection scenario paradigm (Röhner, Degro, Holden, & Schütz, 2025)
- DFG Research Grant: Through a multimethod lens of artificial intelligence and qualitative content analyses: Effects of intelligence and gender on response patterns and elements of the cognitive process of faking (Röhner, 2025)

# References


- Agosta, S., Ghirardi, V., Zogmaister, C., Castiello, U., & Sartori, G. (2011). Detecting fakers of the autobiographical IAT. *Applied Cognitive Psychology, 25*, 299–306. https://doi.org/10.1002/acp.1691
- Ambwani, S., Boeka, A. G., Brown, J. D., Byrne, T. K., Budak, A. R., Sarwer, D. B. et al. (2013). Socially desirable responding by bariatric surgery candidates during psychological assessment. *Surgery for Obesity and Related Diseases, 9* (2), 300–305. https://doi.org/10.1016/j.soard.2011.06.019
- Arthur, W., Jr., Hagen, E. & George, F., Jr. (2021). The lazy or dishonest respondent: Detection and prevention. *Annual Review of Organizational Psychology and Organizational Behavior, 8* (1), 105–137. https://doi.org/10.1146/annurev-orgpsych-0124 20-055324
- Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of personality and social psychology, 97*(3), 533–548. https://doi.org/10.1037/a0016229
- Bäckström, M. & Björklund, F. (2013). Social desirability in personality inventories: Symptoms, diagnosis and prescribed cure. *Scandinavian Journal of Psychology, 54* (2), 152–159. https://doi.org/10.1 111/sjop.12015
- Bäckström, M. & Björklund, F. (2020). The properties and utility of less evaluative personality scales: Reduction of social desirability; increase of construct and discriminant validity. *Frontiers in Psychology, 11*, 560271. https://doi.org/10.3389/fpsyg.2020.560271

# References


- Bäckström, M., Björklund, F. & Larsson, M. R. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43* (3), 335–344. https://doi.org/10.1016/j.jrp.2008.12.013
- Beauducel, A. & Leue, A. (2014). Definitionen, Aufgaben und Rahmenbedingungen. In A. Beauducel & A. Leue (Hrsg.), *Psychologische Diagnostik* (S. 14–22). Göttingen: Hogrefe.
- Birkeland, S. A., Manson, T. M., Kisamore, J. L., Brannick, M. T. & Smith, M. A. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment, 14* (4), 317–335. https://doi.org/10.1111/j.1468-2389.2006.00354.x
- Blasberg, S. A., Rogers, K. H. & Paulhus, D. L. (2014). The Bidimensional Impression Management Index (BIMI): Measuring agentic and communal forms of impression management. *Journal of Personality Assessment*, *96*, 523–531.
- Bressan, M., Rosseel, Y. & Lombardi, L. (2018). The effect of faking on the correlation between two ordinal variables: Some population and Monte Carlo results. *Frontiers in Psychology, 9*, 1876. https://doi.org/10.3389/fpsyg.2018.01876
- Burns, G. N., Shoda, E. A. & Roebke, M. A. (2017). Putting applicant faking effects on personality tests into context. *Journal of Managerial Psychology, 32*, 460–468. https://doi.org/10.1108/JMP-01-2017-0031

# References


- Connelly, B. S. & Chang, L. (2016). A meta-analytic multitrait multirater separation of substance and style in social desirability scales. *Journal of Personality, 84* (3), 319–334. https://doi.org/10.1111/jopy.12161
- Corneille, O., & Hütter, M. (2020). Implicit? What Do You Mean? A Comprehensive Review of the Delusive Implicitness Construct in Attitude Research. Personality and Social Psychology Review, 24(3), 212-232. https://doi.org/10.1177/1088868320911325
- Corneille, O., & Lush, P. (2023). Sixty Years After Orne's American Psychologist Article: A Conceptual Framework for Subjective Experiences Elicited by Demand Characteristics. Personality and Social Psychology Review, 27(1), 83-101. https://doi.org/10.1177/10888683221104368
- Cvencek, D., Greenwald, A. G., Brown, A. S., Gray, N. S., & Snowden, R. J. (2010). Faking of the Implicit Association Test is statistically detectable and partly correctable. *Basic and Applied Social Psychology, 32*, 302–314.
- DeHouwer, J. (2006). What are implicit measures and why are we using them? In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of Implicit Cognition and Addiction* (pp. 11–28). Thousand Oaks: Sage Publications.
- DeSimone, J. A., DeSimone, A. J., Harms, P. D. & Wood, D. (2018). The differential impacts of two forms of insufficient effort responding. *Applied Psychology, 67*, 309–338. https://doi.org/10.1111/apps.12117
- Dilchert, S. & Ones, D. S. (2011). Application of preventive strategies. In M. Ziegler, C. MacCann & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 177–200). New York: Oxford University Press.

# References


- Donovan, J. J., Dwight, S. A. & Hurtz, G. M. (2003). An assessment of the prevalence, severity, and verifiability of entry-level applicant faking using the randomized response technique. *Human Performance, 16* (1), 81–106. https://doi.org/10.1207/S15327043HUP1601_4
- Dunlop, P. D., Holtrop, D., Ashby, L. M., Bharadwaj, A. & Donovan, J. J. (2022). Valence, instrumentality, expectancy, and ability as determinants of faking, and the effects of faking on criterion-related validity. *Journal of Business and Psychology, 37* (6), 1215–1233. https://doi.org/10.1007/s10869-022-09797-0
- Edwards, J. R. (2019). Response invalidity in empirical research: Causes, detection, and remedies. *Journal of Operations Management, 65*, 62–76. https://doi.org/10.1016/j.jom.2018.12.002
- Ellingson, J. E., Sackett, P. R. & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84* (2), 155–166. https://doi.org/10.1037/0021-9010.84.2.155
- Feeney, J. R., Goffin, R. D. & Beshai, S. (2023). Applicant faking warnings: Are they really effective? *Personality and Individual Differences, 200*, 111899. https://doi.org/10.1016/j.paid.2022.1118 99
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology, 27*(4), 307–316. https://doi.org/10.1207/s15324834basp2704_3
- Goldammer, P., Annen, H., Stöckli, P. L. & Jonas, K. (2020). Careless responding in questionnaire measures: Detection, impact, and remedies. *The Leadership Quarterly, 31* (4), 101384. https://doi.org/10.1016/j.leaqua.2020.101384

# References


- Gough, H. G. & Bradley, P. (1996). *California Personality Inventory Manual*. Palo Alto, CA: Consulting Psychologists Press.
- Greenwald, A., McGhee, D. & Schwartz, J. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464
- Griffith, R. L., Chmielowski, T. & Yoshita, Y. (2007). Do applicants fake? An examination of the frequency of applicant faking behavior. *Personnel Review, 36* (3), 341–355. https://doi.org/10.1108/004 83480710731310
- Hildebrand, M., Wibbelink, C. J. M. & Verschuere, B. (2018). Do impression management and self-deception distort self-report measures with content of dynamic risk factors in offender samples? A meta-analytic review. *International Journal of Law and Psychiatry, 58*, 157–170. https://doi.org/10.1 016/j.ijlp.2018.02.013
- Holden, R. R. & Book, A. S. (2011). Faking does distort self-report personality assessment. In M. Ziegler, C. MacCann & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 71–86). Oxford: University Press.
- Huang, J. L., Liu, M. & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology, 100* (3), 828–845. https://doi.org/10.103 7/a0038510
- Jackson, D. N., Ashton, M. C. & Tomes, J. L. (1996). The six-factor model of personality: Facets from the Big Five. *Personality and Individual Differences, 21* (3), 391–402. https://doi.org/10.1016/0191-8869(96)00046-3

# References


- Jaso, B. A., Kraus, N. I. & Heller, A. S. (2021). Identification of careless responding in ecological momentary assessment research: From posthoc analyses to real-time data monitoring. *Psychological Methods*. Manuscript submitted for online publication. https://doi.org/10.1037/met0000312
- Johnson, J. A. (2005). Ascertaining the validity of individual protocols from web-based personality inventories. *Journal of Research in Personality, 39*, 103–129. https://doi.org/10.1016/j.jrp.2004.09.009
- Kallio Strand, K., Bäckström, M. & Björklund, F. (2021). Accounting for the evaluative factor in self-ratings provides a more accurate estimate of the relationship between personality traits and well-being. *Journal of Research in Personality, 93*, 104120.
- Law, S. J., Bourdage, J. & O'Neill, T. A. (2016). To fake or not to fake: Antecedents to interview faking, warning instructions, and its impact on applicant reactions. *Frontiers in Psychology, 7*, 1771. https://doi.org/10.3389/fpsyg.2016.01771
- Li, H., Fan, J., Zhao, G., Wang, M., Zheng, L., Meng, H. et al. (2022). The role of emotions as mechanisms of mid-test warning messages during personality testing: A field experiment. *Journal of Applied Psychology, 107* (1), 40–59. https://doi.org/10.1037/apl0000885
- MacCann, C., Ziegler, M. & Roberts, R. D. (2011). Faking in personality assessment: Reflections and recommendations. In M. Ziegler, C. MacCann & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 309–329). Oxford: University Press.

# References


- Mancosu, M., Ladini, R. & Vezzoni, C. (2019). “Short is better”. Evaluating the attentiveness of online respondents through screener questions in a real survey environment. *Bulletin of Sociological Methodology, 141* (1), 30–45.
- Maniaci, M. R. & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality, 48*, 61–83. https://doi.org/10.1016/j.jrp.2013. 09.008
- Moosbrugger, H. & Brandt, H. (2020a). Antwortformate und Itemtypen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 91–117). Berlin: Springer. https://doi.org/10.1007/978-3-662-61532-4_5
- Moosbrugger, H. & Brandt, H. (2020b). Itemkonstruktion und Antwortverhalten. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 67–89). Berlin: Springer. https://doi.org/10.1007/978-3-662-61532-4_4
- Musch, J. (1999). *Ehrliche Antworten auf peinliche Fragen: Die Randomized-Response-Technik*. Beitrag der German Online Research Tagung.
- Paulhus, D. L. (1994). *Balanced Inventory of Desirable Responding: Reference manual for BIDR version 6*. Unpublished manuscript, University of British Columbia, Vancouver, Canada.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

# References


- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Paunonen, S. V. & LeBel, E. P. (2012). Socially desirable responding and its elusive effects on the validity of personality assessments. *Journal of Personality and Social Psychology, 103* (1), 158–175. https://doi.org/10. 1037/a0028165
- Pella, R. D., Hill, B. D., Shelton, J. T., Elliott, E. & Gouvier, W. D. (2012). Evaluation of embedded malingering indices in a non-litigating clinical sample using control, clinical, and derived groups. *Archives of clinical neuropsychology: The official journal of the National Academy of Neuropsychologists, 27* (1), 45–57. https://doi.org/10.1093/arclin/acr090
- Rios, J. A., Guo, H., Mao, L. & Liu, O. L. (2017). Evaluating the Impact of careless responding on aggregated-scores: To filter unmotivated examinees or not? *International Journal of Testing, 17* (1), 74–104. https://doi.org/10.1080/15305058.2016.1231193
- Röhner, J., Degro, M., Holden, R. R., & Schütz, A. (2025). A registered report to disentangle the effects of frame of reference and faking in the personnel-selection scenario paradigm. *International Journal of Selection and Assessment* [Manuscript in Revision]
- Röhner, J., Holden, R. R. & Schütz, A. (2023). IAT faking indices revisited: Aspects of replicability and differential validity. *Behavior Research Methods, 55*, 670–693. https://doi.org/10.3758/s13428-02 2-01845-0

# References


- Röhner, J., & Iliescu, D. (2023). On the death of Implicit Association Tests (IATs) [Editorial]. *European Journal of Psychological Assessment*, *39*(5), 317–322. https://doi.org/10.1027/1015-5759/a000778

- Röhner, J., & Lai, C. K. (2021). A diffusion model approach for understanding the impact of 17 interventions on the race Implicit Association Test. *Personality and Social Psychology Bulletin, 47*(9), 1374-1389. https://doi.org/10.1177/0146167220974489

- Röhner, J., Schröder-Abé, M. & Schütz, A. (2011). Exaggeration is harder than understatement, but practice makes perfect! *Experimental Psychology, 58*, 464–472. https://doi.org/10.1027/1618-3169/a000114

- Röhner, J., Schröder-Abé, M., & Schütz, A. (2013). What do fakers actually do to fake the IAT? An investigation of faking strategies under different faking conditions. *Journal of Research in Personality, 47*(4), 330–338. https://doi.org/10.1016/j.jrp.2013.02.009

- Röhner, J., Thoss, P. J. & Schütz, A. (2022). Lying on the dissection table: Anatomizing faked responses. *Behavior Research Methods, 54*, 2878–2904. https://doi.org/10.3758/s13428-021-01770-8

- Röhner, J., & Schütz, A. (2019b). Faking behavior. In V. Zeigler-Hill and T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences*. Springer. https://doi.org/10.1007/978-3-319-28099-8_2341-1

# References


- Röhner, J., & Schütz, A. (2022). Psychologische Diagnostik [Psychological assessment]. In A. Schütz, M. Brand, H. Selg, & S. Lautenbacher (Eds.), Psychologie: Eine Einführung in ihre Grundlagen und Anwendungsfelder [Psychology: An introduction to its foundations and fields of application] (6th ed., pp. 199–218). Kohlhammer.

- Röhner, J., & Schütz, A. (2025). Phänomene der Antwortverzerrung in der Psychologischen Diagnostik [Response biases in psychological assessment]. Hogrefe. https://doi.org/10.1024/85915-000

- Röhner, J., Schütz, A., & Ziegler, M. (2025). Faking in self-report personality scales: A qualitative analysis and taxonomy of the behaviors that constitute faking strategies. *International Journal of Selection and Assessment* [Manuscript in Revision]

- Salgado, J. F. (2016). A theoretical model of psychometric effects of faking on assessment procedures: Empirical findings and implications for personality at work. *International Journal of Selection and Assessment, 24* (3), 209–228. https://doi.org/10.1111/ijsa.12142

- Schroeders, U., Schmidt, C. & Gnambs, T. (2022). Detecting careless responding in survey data using stochastic gradient boosting. *Educational and Psychological Measurement, 82* (1), 29–56. https://doi.org/10.1177/00131644211004708

- Van Iddekinge, C. H., Raymark, P. H. & Roth, P. L. (2005). Assessing personality with a structured employment interview: Construct-related validity and susceptibility to response inflation. *Journal of Applied Psychology, 90*, 536–552. https://doi.org/10.1037/0021-9010.90.3.536

# References


- Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28*, 186–191. https://doi.org/10.1007/s10862-005-9004-7
- Youngjohn, J., Lees-Haley, P. & Binder, L. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists, 14*, 511–515.
- Ziegler, M. & Buehner, M. (2009). Modeling socially desirable responding and its effects. *Educational and Psychological Measurement, 69* (4), 548–565. https://doi.org/10.1177/0013164408324469
- Ziegler, M., MacCann, C. & Roberts, R. D. (2011). Faking: Knowns, unknowns, and points of contention. In M. Ziegler, C. MacCann & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (S. 3–16). Oxford: University Press.

# 6. Discussion