# *Surely you're sure, but are you right?*

# A Three-day Randomized Controlled Trial on the Confidence-Accuracy Relationship

Sophie Bonsack, Leni Engelhardt, Rasmus Glück, Carolina Lange, Elias Thelemann  Leitung: Dr. Jürgen M. Kaufmann

## Introduction

### Theoretical Backround

- Eyewitness identifications are **often inaccurate**, making confidence judgments important in legal settings (1).
- Recent studies show that **confidence** (the belief that one's judgement is correct) is positively related to identification **accuracy** (the actual correctness), however, this relationship is not yet fully understood (2).
- Our **previous study** indicates that the relationship might be robust to **personality differences.**
- While the accuracy appears to be invariant to training (1), metacognitive skills can be improved through feedback (3).
- **Research question**: Can individualized feedback on accuracy and confidence enhance the confidence–accuracy relationship?

### Hypotheses

***H1*** Higher levels of confidence predict higher odds of completing the line-up correctly (accurately) in all conditions.

***H2*** Intervention effect: The confidence–accuracy relationship increases in the experimental group (i.e., higher correlation in post-test as compared to pre-test and the control group)

***H3*** Accuracy does not increase in both groups (i.e., in post-test as compared to pre-test)

## Methods

### Participants

N = 103 (EG = 55; CG = 48)  ♀ 80  ♂ 21  divers 2

Age: 18-80 [*M* = 27.2, *SD* = 13.8]

### Stimuli

- 30 line-ups with 8 phenotypically similar neutral faces (Chicago Face Database)
- Line-ups either included a learned face (target present) or only unknown faces (target absent), counterbalanced
- Line-ups covered individuals from diverse ethnic backgrounds


*Figure 1.* Example of a sequential lineup


*Figure 2.* Example of a simultaneous lineup

### Analytical Strategy

Multilevel logistic regression analyses

- Level 2: participants; Level 1: trials (identification tasks)
- *ICC* = 0.08, *95% CI* [0.05, 0.12]
- Outcome: trial-wise accuracy (dummy-coded)

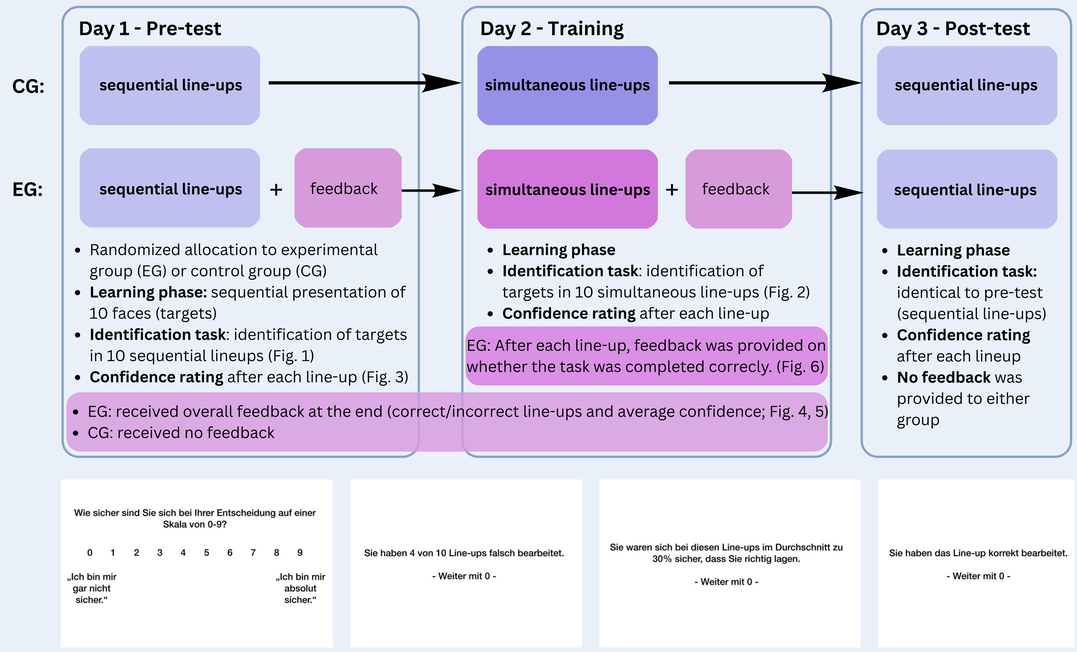





*Figure 3.* Confidence-rating




*Figure 4.* Example of the overall feedback




*Figure 5.* Example of the overall feedback




*Figure 6.* Example of the direct feedback

## Results

***H1*** ✔

**Higher levels of confidence** predicted **higher odds of accuracy**, both

a) **participant-wise** (confidence group-mean [i.e., context variable]), *B* = 0.24, *SE* = 0.04, *p* < .001, *OR* = 1.27, *95% CI* [1.17, 1.39].

b) **trial-wise** at **pre-test** (confidence [group-mean centered]), *B* = 0.15, *SE* = 0.05, *p* = .004, *OR* = 1.17, *95% CI* [1.05, 1.30].

c) **trial-wise** at **post-test** (no significant difference to pre-test, i.e., interaction of confidence [group-mean centered] x test phase), *B* = 0.11, *SE* = 0.08, *p* = .177, *OR* = 1.11, *95% CI* [0.96, 1.29].

***H2*** ✘

**Absence of intervention effect** (three-way interaction of confidence [group-mean centered] x group membership x test phase), *B* = -0.09, *SE* = 0.10, *p* = .373, *OR* = 0.91, *95% CI* [0.74, 1.12].

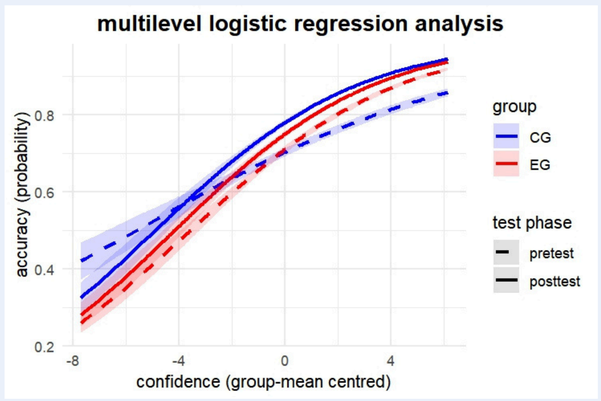


*Figure 7.* Multilevel logistic regression analyses. Model fit: $\Delta\chi^2(8) = 136.92$, p < 0.001, Random intercept: $\sigma^2 = 0.20$, *SD* = 0.45

***H3*** ✘

**Accuracy increased in both groups** (coefficients for post-test)

a) CG (test phase), *B* = 0.40, *SE* = 0.15, *p* = .009, *OR* = 1.50, *95% CI* [1.10, 2.03].

b) EG (no significant difference to CG, i.e., interaction of group membership x test phase), *B* = -0.20, *SE* = 0.21, *p* = .341, *OR* = 0.82, *95% CI* [0.54, 1.24]; although increase was descriptively smaller than in CG.

Exploratory analyses revealed significant differences in target present vs. target absent lineups in regard to the confidence–accuracy relationship.

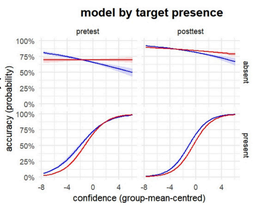


*Figure 8.* Model by target presence

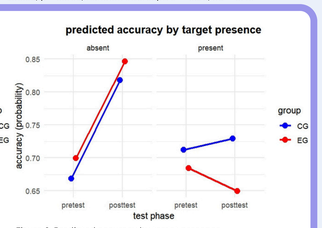


*Figure 9.* Predicted accuracy by target presence

## Discussion

***H1***

- Corresponding with our previous study and the current state in research (2), **confidence was a relevant predictor for the accuracy** of eyewitness identifications.
- **Both trial-wise** and **participant-wise**, higher confidence ratings predicted higher odds for completing the line-up correctly (i.e., accurately).
- This contradicts the opinions often held by both experts and policy-makers for the past decades (4).

***H2***

- **No evidence** for the postulation that the confidence-accuracy relationship might be enhanced through individualized feedback (i.e., **intervention effect**; which has not been examined by any prior study) was found in our sample.
- However, this result found in our specific sample might be due the limited sample size in our study, or the short training period compared to similar meta-cognitive training studies (3).

***H3***

- In contrast to other empirical findings (1), we observed a significant **training effect on accuracy** in our sample.
- Further analyses indicate that this increase of accuracy occurs due to more overall line-up rejections and because more correct line-up rejections are made specifically in **target absent** trials.
- According to signal detection theory, which is widely established in eyewitness research (5), this effect may be explained by a **bias shift** towards more conservative decisions resulting from application of different decision criteria.
- There was **no positive confidence–accuracy relationship in target absent line-ups**, indicating that choosing a target and rejecting a line-up are distinct mechanisms that should be considered separately.

### References

1. Towler, A., Kemp, R. I., Burton, A. M., Dunn, J. D., Wayne, T., Moreton, R., & White, D. (2019). Do professional facial image comparison training courses work?. *PloS one, 14*(2), e0211037. https://doi.org/10.1371/journal.pone.0211037
2. Seale-Carlisle, T. M., Quigley-McBride, A., Teitcher, J. E. F., Crozier, W. E., Dodson, C. S. & Garrett, B. L. (2024). New Insights on Expert Opinion About Eyewitness Memory Research. *Perspectives On Psychological Science, 20*(5), 903–924. https://doi.org/10.1177/17456916241234837
3. Carpenter, J., Sherman, M. T., Kievit, R. A., Seth, A. K., Lau, H., & Fleming, S. M. (2019). Domain-general enhancements of metacognitive ability] through adaptive training. *Journal of Experimental Psychology. General, 148*(1), 51–64. https://doi.org/10.1037/xge0000505
4. Wixted, J. T., & Wells, G. L. (2017). The Relationship Between Eyewitness Confidence and Identification Accuracy: A New Synthesis. *Psychological Science in the Public Interest, 18*(1), 10-65. https://doi.org/10.1177/1529100616686966
5. Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review, 121*(2), 262–276. https://doi.org/10.1037/a0035940
6. Stoet G. (2010). PsyToolkit: a software package for programming psychological experiments using Linux. *Behavior research methods, 42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096
7. Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology, 44*(1), 24-31.

**Ethics Approval:** This research project received positive ethics approval (Approval no. FSV 26/110).