

Indicators for the ex post detection of faking in survey data constructed from responses to the Big Five Inventory-10

Theoretical Background

Surveys are an essential data collection method for Social Science research.

Faking of face-to-face interviews by interviewers (fakers) threatens the quality of survey data and the validity of conclusions drawn from the data.

Survey participants and fakers of survey questionnaires differ in their motivational processes when responding to questionnaire items (Menold, Storfinger, & Winker, 2011).

- Participants want to complete the questionnaire with minimal effort.
- Fakers try to stay undetected when faking the questionnaire.

Hypothesis: Due to these differences, response sets and satisficing differentially impact participant's and fakers' responses to questionnaire items.

Goal: Which indicators of response sets and satisficing behaviour are sensitive to faking and can be used to identify faked interviews?

Method

A subset of 101 interviews of German ISSP 2008 participants (51.5% male; Age $M=53.5$, $SD=15.9$) were faked by providing fakers with 11 sociodemographic characteristics, e.g. sex, age, education, party preference etc.

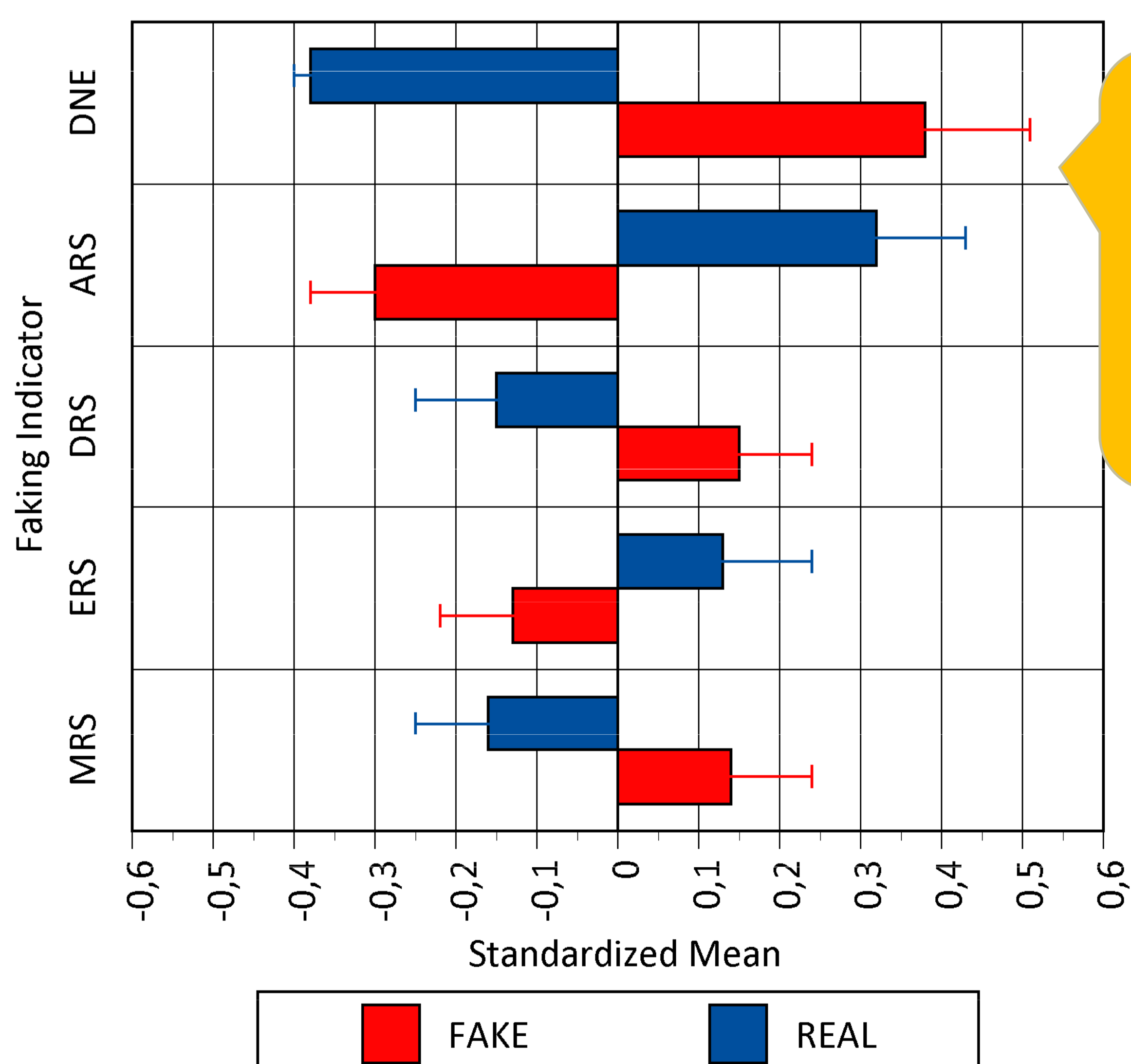
Measure: Ultra-short 10 item measure of the Big Five, BFI-10 (Rammstedt & John, 2007).

Faking indicators

DNE	ARS	DRS	ERS	MRS
Don't know category endorsement	Acquiescent response style	Disacquiescent response style	Extreme response style	Middle response style

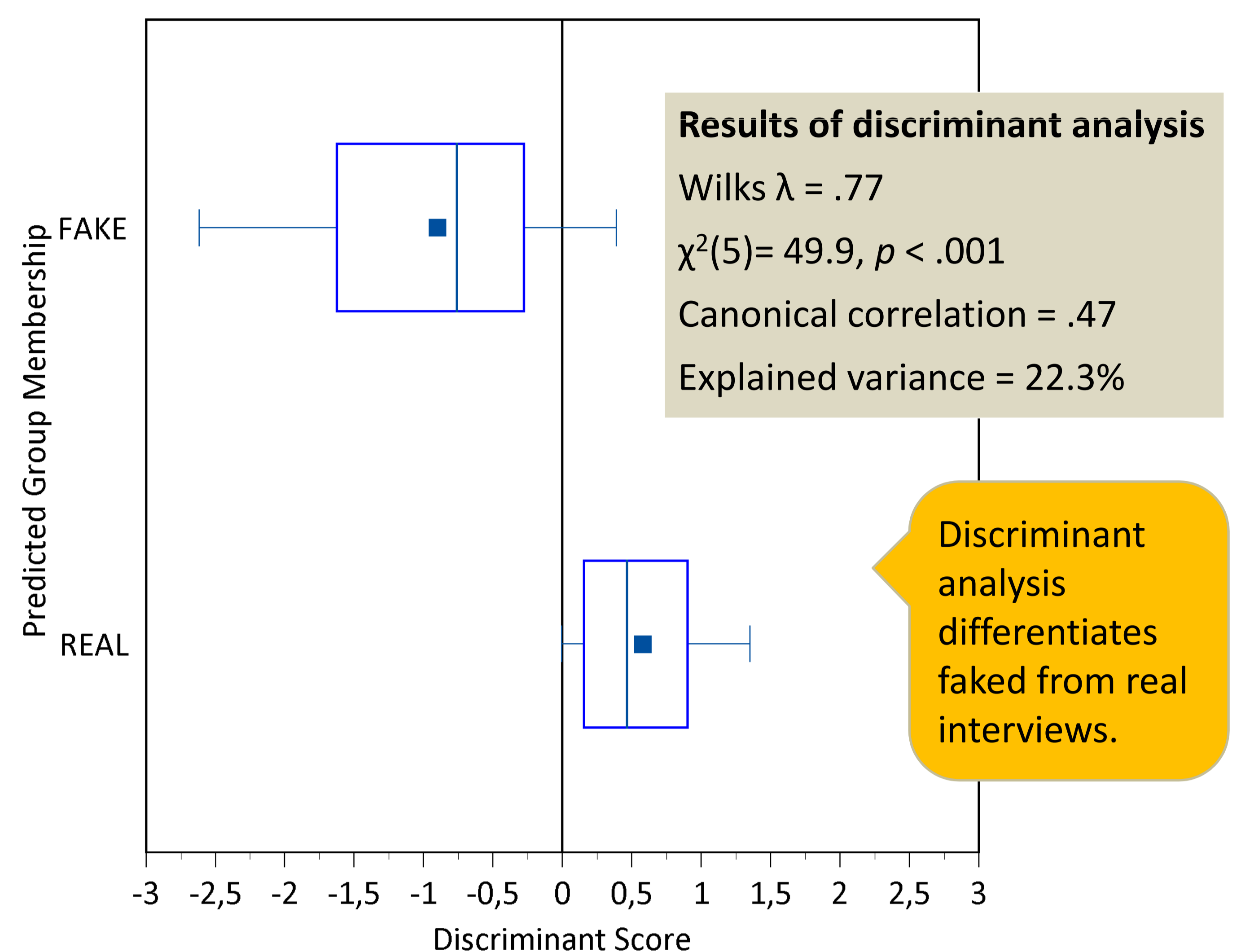
Results

Which faking indicators are useful to differentiate faked from real interviews?



Strongest differences between real and faked interviews were found for DNE and ARS

Can faking indicators reliably differentiate faked from real interviews?



Faking indicators	Discriminant loadings
DNE	-.77
ARS	.62
DRS	-.31
ERS	.25
MRS	-.28

Most relevant faking indicators for the discrimination of real and faked interviews are DNE and ARS.

73% of faked interviews were correctly identified.

How many faked interviews can be identified by the discriminant analysis?

	Predicted group membership		
	Real interviews	Faked interviews	Total
Original group membership	Real interviews (82.2%)	18 (17.8%)	101
	Faked interviews (51.5%)	49 (48.5%)	101

Hit ratio = 65.3%; Positive predictive value (precision) = 73.1%; Negative predictive value = 61.5%

Summary & Conclusion

Substantial differences between real and faked interviews were found especially for DNE and ARS.

A substantial number of faked interviews were identified correctly by using faking indicators calculated from responses to ten items only (BFI-10).

Outlook: Faking indicators can be used with statistical procedures, e.g. cluster analysis, in order to identify "at risk interviewers" in surveys.

Results could be further improved by using more items to construct faking indicators.