

# Response Biases

I. Types of Response Bias

II. Acquiescence

III. Extreme responding

IV. Social desirability

V. Malingering

VI. Careless or random responding

VII. How Common and Problematic are Response Biases?

## I. Types of Response Bias

*Response bias* is a threat to validity of a measure

Scores may partially reflect a tendency to respond in a systematic way rather than the construct

Note that a “bias” is systematic variation not random variation, may have no impact whatsoever reliability

## I. Types of Response Bias

Response biases may occur consistently within an individual (a stable characteristic of the respondent), known as *response style*

Example: personality characteristic to please others

A *response set* involves a temporary systematic response tendency due to testing conditions or other situational factors (not a stable characteristic of the respondent, affects)

Example: aversive, noisy environment causes rushed answers

## II. Acquiescence

*Acquiescence bias* – respondent agrees with the item regardless of the respondent's true attitude

“Yea-saying” and “nay-saying” bias

May inflate estimates of correlations among items worded in the same direction (a method bias)

May exaggerate correlations between constructs if both affected by acquiescence bias

## II. Acquiescence

Most common method used for mitigating acquiescence bias is using reverse-worded items

Fully neutralizing the effects requires equal number and magnitude of items (i.e., *balanced scale*)

Can investigate by examining correlation between positive half and negative half of measure and comparing means of (reverse-scored) negatively worded items to the positively worded items

Confirmatory factor analysis method that uses wording direction as a “method factor”

## II. Acquiescence

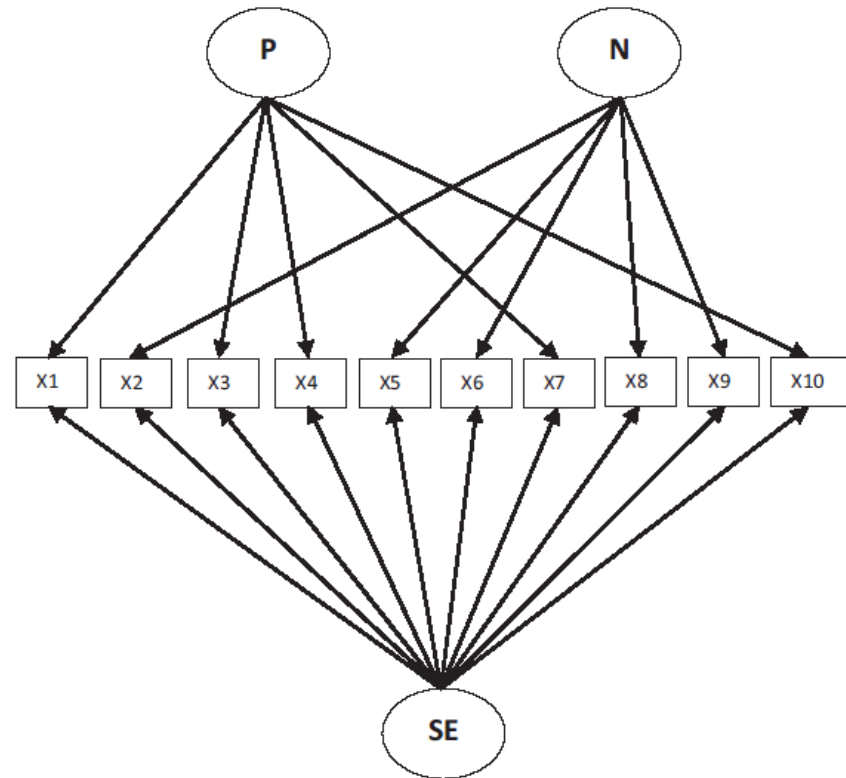


FIGURE 4.—Bifactor model of Rosenberg Self-Esteem Scale. *Note.* P = positive self-esteem; N = negative self-esteem; SE = self-esteem.

McKay, M. T., Boduszek, D., & Harvey, S. A. (2014). The Rosenberg Self-Esteem Scale: a bifactor answer to a two-factor question?. *Journal of personality assessment*, 96(6), 654-660.

### III. Extreme Responding

*Extreme responding* – a tendency to use the extreme ends of the scale

*Moderate responding* – tendency to avoid the extreme ends of the scale

Although individuals are consistent in extreme responding over time, it is not strongly associated with any personality trait (Schwartz & Sudman, 1996)

Difficult to distinguish between extreme responding and extreme attitudes

### III. Extreme Responding

Extreme responding can artificially inflate correlations among items and among different self-report measures



### III. Extreme Responding

Extreme responding might be prevented by more neutral wording, more response scale points, or forced choice (although forced choice has other disadvantages)

## IV. Social Desirability

*Social desirability* – tendency to respond in a way that is the most socially acceptable

Perhaps most studied of all response biases

May be a result of content or context

Examples:

Racial prejudice measures are susceptible “do you believe in treating everyone fairly?”

When respondent has something to gain e.g., job applicant is asked “are you a hard worker?”

## IV. Social Desirability

Two possible forms of social desirability according to Paulhus (2002)

Impression management – attempt to look good in front of others

Self-deception – unrealistically positive view of themselves (e.g., everyone is above average intelligence)

## IV. Social Desirability

As with most other response biases, social desirability has the potential to inflate correlations with other self-report measures

## IV. Social Desirability

Social desirability can be addressed by

- Subtle items (low face validity)

- Nonobvious purpose of the study

- Anonymous responses

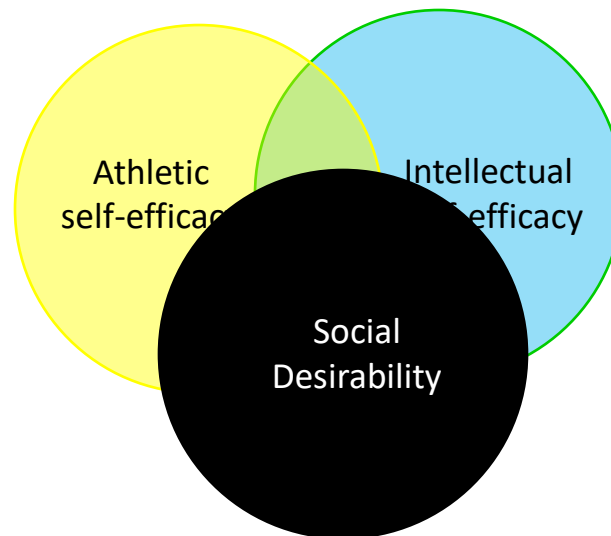
- Use euphemistic or dysphemistic terms (in general, more neutral wording)

- Prior examination of discriminant validity of measure with social desirability scale

- Use of social desirability scale as a control variable

## IV. Social Desirability

Statistically controlling for social desirability



Multiple regression analysis would allow the researcher to examine the association between intellectual efficacy and athletic efficacy that is ***not due*** to social desirability

## IV. Social Desirability

### Social desirability measures

Crowne-Marlowe social desirability scale is most widely known  
(Crowne & Marlowe, 1960)

Balanced Inventory of Desirable Responding (BIDR: Paulus, 1984, 1988)

## IV. Social Desirability

### 16-item short form of BIDR (Hart, Ritchie, Hepper, & Gebauer, 2015)

**Table 2.** Factor Loadings Per Sample for the Two-Factor BIDR-16 Model in Study 1.

Factor	Item	Sample				
		All <sup>a</sup>	Women <sup>b</sup>	Men <sup>c</sup>	UK <sup>d</sup>	USA <sup>e</sup>
SDE	Item 4 “not always honest” (r)	.46	.45	.49	.38	.52
	Item 5 “know why like things”	.42	.43	.34	.37	.41
	Item 10 “hard to shut off a disturbing thought”(r)	.40	.37	.43	.32	.47
	Item 11 “never regret decisions”	.56	.56	.55	.53	.53
	Item 12 “can’t make up my mind” (r)	.44	.40	.51	.41	.48
	Item 15 “completely rational”	.46	.47	.41	.47	.39
	Item 17 “confident in judgements”	.52	.55	.39	.58	.41
	Item 18 “doubted ability as a lover” (r)	.33	.31	.36	.32	.41
IM	Item 21 “sometimes tell lies” (r)	.56	.56	.57	.59	.55
	Item 22 “never cover up mistakes”	.47	.46	.53	.44	.46
	Item 23 “taken advantage of someone”(r)	.48	.46	.58	.48	.53
	Item 25 “sometimes try to get even” (r)	.44	.44	.40	.39	.49
	Item 27 “said something bad about a friend” (r)	.57	.60	.49	.52	.61
	Item 28 “avoid listening”	.48	.53	.33	.47	.50
	Item 36 “never take things”	.38	.38	.39	.40	.38
	Item 40 “don’t gossip”	.53	.56	.43	.47	.59

Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open*, 5(4), 1-9.



## V. Malingering

*Malingering* – tendency to exaggerate psychological problems  
 (“faking bad”)

Occurs when it is in the respondent’s best interest to look worse

Examples:

- Criminal defendants who want to use an insanity defense

- Criminal competence (e.g., accused of negligence)

- Lawsuits (e.g., disability or other insurance claim)

## V. Malingering

Malingering may be addressed by

- Subtle questions

- Nonobvious intent of measure

- Embedding bias detection items (e.g., “lie” or “validity” scales)

## V. Malingering

### MMPI-2 contains *validity scales*

L or Lie scale to detect lying or faking (deny common weaknesses)

F scale to detect unusual or random response

Back F scale  $F_b$  unusual or random response at the end of the test  
(prevent careless or random responding)

K to detect faking

## V. Malingering

MMPI-2 L (“Lie”) scale examples:

“I do not always tell the truth”

“Once in a while I think about things too bad to talk about”

“My table manners are not quite as good at home as when I am out in company”

## V. Malingering

Evidence is mixed on the impact of malingering/faking on personality, hiring, and clinical measurement and whether validity scales are effective and identifying it (Hahn, 2005; Kroger & Turnbull, 1975; Piedmont et al., 2000; Ross et al., 1998; Worthington & Schottman, 1986; Zicker & Robie, 1999)

Possible to embed one or two items as a check, particularly when dishonesty might be expected or particularly consequential

## VI. Careless or Random Responding

*Careless or random responding* – inconsistent responding due to lack of motivation

More likely when incentives are low, choice is low, and burden is high

May more likely occur at the end of a long measure

This pattern would suggest no systematic bias overall

## VI. Careless or Random Responding

### Steps that will help prevent careless responding

- Improve incentives
- Stress importance of research
- Increase perceived choice in participation
- Keep overall survey length as short as possible to reduce fatigue

## VII. How Common and Problematic are Response Biases?

No simple or definite answer—most likely depends on topic, context, motivations, incentives, and individual differences

Few good studies on prevalence, most studies on social desirability and effects of faking or detection of faking

Debate and evidence exists on both sides about how problematic response biases are (e.g., Gove, McCorkle, Fain, & Hughes, 1976; Nederhoff, 1985)



## VII. How Common and Problematic are Response Biases?

Scale validation studies can and often do demonstrate discriminant validity with social desirability scales

Some biases such as acquiescence bias can be easily reduced by inclusion of reverse-worded items (a fairly common practice) or careful consideration of item wording

Careless and random responding pattern would impact reliability, and if scale has high reliability, it would suggest this type of responding is rare or has minimal impact

Studies can and often do consider response biases as a potential confound and attempt to address through design, measurement, or analyses

## VII. How Common and Problematic are Response Biases?

Some studies (e.g., Hogan, Barrett, & Hogan, 2007) demonstrate little impact of response bias or faking on conclusions of the study

For example, Hogan and colleagues showed that applicants did not show much additional effort to manage their impressions when reapplying for a job after being denied

Validation studies for many phenomena have been conducted to support reasonable accuracy of self-report, especially under appropriate measurement conditions (e.g., report of health conditions; depression; physical activity)