Psychometric Barriers: Response Biases

PSYC3302: Psychological Measurement and Its Applications

Mark Hurlstone
Univeristy of Western Australia

Week 7
Learning Objectives

• Introduce problem of response biases:
  • Acquiescence Bias
  • Extreme and Moderate Responding
  • Social Desirability
  • Malingering
  • Random Responding
  • Guessing

• Methods that have been used to detect the problem
• Methods that have been used to deal with the problem
Response Bias

Imagine that you have taken a personality test as part of a job interview and you were asked to respond to these questions:

1. "Have you ever stolen anything from an employer?"
2. "Do you always tell the truth to others?"

The true answer to 1 may be "yes" and the true answer to 2 may be "no."

But your actual responses might be biased by your desire to impress the employer.

Such a bias would affect the validity of interpretations of your questionnaire responses as reflecting honesty or integrity.
Response Bias

• We want to interpret responses on psychological measures as being an accurate reflection of individuals’ true psychological characteristics

• However, responses to psychological measures can be systematically biased for a variety of reasons

• These biases are important because they can harm the psychometric quality of psychological tests

• Specifically, they can diminish test reliability and the validity of test score interpretations

• Response biases can therefore damage our ability to use psychological measures in a meaningful way
Types of Response Biases

- There are various different types of response biases
- They can be affected by:
  - the testing content
  - the testing format
  - the testing context
  - respondent’s conscious efforts to distort "reality"
  - unconscious factors that bias responses
Types of Response Biases

• The response biases that will be discussed in this lecture are:
  • Acquiescence Bias
  • Extreme and Moderate Responding
  • Social Desirability
  • Malingering
  • Random Responding
  • Guessing
The response biases that will be discussed in this lecture are:

- Acquiescence Bias
- Extreme and Moderate Responding
- Social Desirability
- Malingering
- Random Responding
- Guessing
Acquiescence Bias

- Acquiescence bias occurs when an individual agrees with statements without regard for the meaning of those statements.

- For example, an inventory may have items such as:
  - "I enjoy my job"
  - "I dislike my job"

- In its extreme form, people who engage in acquiescence will respond "strongly agree" to both of these items even though they are polar opposites of the same construct.

- These are people who will agree or say "yes" to just about everything.
Acquiescence Bias

- Imagine that an organisational psychologist is interested in the association between job satisfaction and perceived prestige.
- Suppose job satisfaction is measured with the following items:
  1. "I really enjoy my work"
  2. "I find my work personally fulfilling"
  3. "In general, I am satisfied with the day to day aspects of my job"
  4. "There is very little I would change about my job"
- Response are made on a 7-point scale (1 = strongly disagree; 7 = strongly agree)
- Scores on a each item are simply summed to give an overall measure of job satisfaction.
Acquiescence Bias

- The items’ phrasing is an important issue in the job satisfaction scale
- Note that all items are "positively keyed"—a positive (agreement) response to each item reflects a relatively high level of the construct (job satisfaction) being assessed
- This renders the questionnaire susceptible to the effects of an acquiescence response bias
- Consider a hypothetical "all-seeing" example in which we know which respondents exhibited an acquiescence bias
# Acquiescence Bias

(a) Responses to original tests

<table>
<thead>
<tr>
<th>Participant</th>
<th>Acquiescence</th>
<th>JS Items</th>
<th>PP Items</th>
<th>Valid Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>JS Total</td>
<td>PP Total</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Y</td>
<td>6 5 7 6 24</td>
<td>5 5 4 5 19</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>7 5 6 7 25</td>
<td>2 2 2 1 7</td>
<td>25 7</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>3 4 5 4 16</td>
<td>5 4 5 4 18</td>
<td>16 18</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>6 6 6 7 25</td>
<td>5 5 5 5 20</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>N</td>
<td>1 4 2 3 10</td>
<td>1 2 1 2 6</td>
<td>10 6</td>
</tr>
<tr>
<td>6</td>
<td>N</td>
<td>3 2 4 3 12</td>
<td>3 3 3 3 12</td>
<td>12 12</td>
</tr>
</tbody>
</table>

Correlation between JS and PP = –.43

JS = job satisfaction; PP = perceived prestige
Acquiescence Bias

- In this "all-seeing" example, we know that participants 1 and 4 were acquiescent during responding.
- Both participants agreed strongly with all four items even though they honestly might not be satisfied with their jobs.
- By contrast, participant 2 also agreed with all four items because she is genuinely satisfied with her job.
- In practice, we would not be able to distinguish the acquiescent responders from the valid responders.
Acquiescence Bias

### Types of Response Biases
- Acquiescence Bias
- Extreme and Moderate Responding
- Social Desirability
- Malingering
- Random Responding
- Guessing

### Methods For Coping With Response Bias
- Managing Testing Context
- Managing Test Content
- Specialised Tests

#### (a) Responses to original tests

<table>
<thead>
<tr>
<th>Participant</th>
<th>Acquiescence</th>
<th>JS Items</th>
<th>PP Items</th>
<th>Valid Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>6 5 7 6 24</td>
<td>5 5 4 5 19</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>7 5 6 7 25</td>
<td>2 2 2 1 7</td>
<td>25 7</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>3 4 5 4 16</td>
<td>5 4 5 4 18</td>
<td>16 18</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>6 6 6 7 25</td>
<td>5 5 5 5 20</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>N</td>
<td>1 4 2 3 10</td>
<td>1 2 1 2 6</td>
<td>10 6</td>
</tr>
<tr>
<td>6</td>
<td>N</td>
<td>3 2 4 3 12</td>
<td>3 3 3 3 12</td>
<td>12 12</td>
</tr>
</tbody>
</table>

Correlation between JS and PP = −.43

JS = job satisfaction; PP = perceived prestige
Acquiescence Bias

- In this "all-seeing" example, we know that participants 1 and 4 were acquiescent during responding
- Both participants agreed strongly with all four items even though they honestly might not be satisfied with their jobs
- By contrast, participant 2 also agreed with all four items because she is genuinely satisfied with her job
- In practice, we would not be able to distinguish the acquiescent responders from the valid responders
Acquiescence Bias: Interpreting Individual Test Scores

- Acquiescence bias can make it difficult to use test scores to identify individuals who have a high level of the construct being assessed.
- Imagine, for example, that a human resources director administers a Conscientiousness scale to a group of job applicants.
- Suppose that he is unaware of the problem of acquiescence bias.
- He could end up employing an acquiescent responder, instead of someone who is a truly conscientious worker.
Acquiescence Bias: Spurious Correlations

- If multiple tests are "contaminated" by acquiescence bias the tests will be more strongly correlated than the underlying constructs.
- If respondents are acquiescent on one test they are likely to be so on another test, giving rise to relatively high scores on both.
- Returning to the job satisfaction example, suppose that the organisational psychologist also administered a four-item measure of perceived prestige answered on a 5-point scale (1 = strongly disagree; 5 = strongly agree).
- Suppose also that participants 1 and 4 once again exhibit an acquiescence bias.
Acquiescence Bias: Spurious Correlations

- Across all six participants, the correlation between job satisfaction and perceived prestige is $r = .43$
- However, in our "all-seeing" state, we can also compute the correlation between the two measures for the four participants who responded validly
- The valid responder correlation is much weaker at $r = -.09$
- In practice, the researcher only has access to the original correlation—she can’t distinguish "valid" from "invalid" responders
- Thus, she would incorrectly conclude that there is a link between the two constructs
Acquiescence Bias: Spurious Correlations

- Across all six participants, the correlation between job satisfaction and perceived prestige is $r = .43$.
- However, in our "all-seeing" state, we can also compute the correlation between the two measures for the four participants who responded validly.
- The valid responder correlation is much weaker at $r = -0.09$.
- In practice, the researcher only has access to the original correlation—she can’t distinguish "valid" from "invalid" responders.
- Thus, she would incorrectly conclude that there is a link between the two constructs.
## Acquiescence Bias

<table>
<thead>
<tr>
<th>Participant</th>
<th>Acquiescence</th>
<th>JS Items 1</th>
<th>JS Items 2</th>
<th>JS Items 3</th>
<th>JS Items 4</th>
<th>JS Total</th>
<th>PP Items 1</th>
<th>PP Items 2</th>
<th>PP Items 3</th>
<th>PP Items 4</th>
<th>PP Total</th>
<th>Valid Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>24</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>19</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>25</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>16</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>25</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>20</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>N</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>10</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>N</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

Correlation between JS and PP = \(-.43\)

\(JS = \text{job satisfaction}; \ PP = \text{perceived prestige}\)
Acquiescence Bias: Spurious Correlations

- Across all six participants, the correlation between job satisfaction and perceived prestige is $r = .43$
- However, in our "all-seeing" state, we can also compute the correlation between the two measures for the four participants who responded validly
- The valid responder correlation is much weaker at $r = -.09$
- In practice, the researcher only has access to the original correlation—she can’t distinguish "valid" from "invalid" responders
- Thus, she would incorrectly conclude that there is a link between the two constructs
**Acquiescence Bias: Spurious Correlations**

- Across all six participants, the correlation between job satisfaction and perceived prestige is $r = .43$

- However, in our "all-seeing" state, we can also compute the correlation between the two measures for the *four participants who responded validly*

- The valid responder correlation is much weaker at $r = -.09$

- In practice, the researcher only has access to the original correlation—she can’t distinguish "valid" from "invalid" responders

- Thus, she would incorrectly conclude that there is a link between the two constructs
Acquiescence Bias

<table>
<thead>
<tr>
<th>Participant</th>
<th>Acquiescence</th>
<th>JS Items</th>
<th>PP Items</th>
<th>Valid Responders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>JS Total</td>
<td>PP Total</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Y</td>
<td>6 5 7 6 24</td>
<td>5 5 4 5 19</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>N</td>
<td>7 5 6 7 25</td>
<td>2 2 2 1 7</td>
<td>25 7</td>
</tr>
<tr>
<td>3</td>
<td>N</td>
<td>3 4 5 4 16</td>
<td>5 4 5 4 18</td>
<td>16 18</td>
</tr>
<tr>
<td>4</td>
<td>Y</td>
<td>6 6 6 7 25</td>
<td>5 5 5 5 20</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>N</td>
<td>1 4 2 3 10</td>
<td>1 2 1 2 6</td>
<td>10 6</td>
</tr>
<tr>
<td>6</td>
<td>N</td>
<td>3 2 4 3 12</td>
<td>3 3 3 3 12</td>
<td>12 12</td>
</tr>
</tbody>
</table>

Correlation between JS and PP = -.43

JS = job satisfaction; PP = perceived prestige
Acquiescence Bias: Spurious Correlations

- Across all six participants, the correlation between job satisfaction and perceived prestige is $r = .43$
- However, in our "all-seeing" state, we can also compute the correlation between the two measures for the four participants who responded validly
- The valid responder correlation is much weaker at $r = -.09$
- In practice, the researcher only has access to the original correlation—she can’t distinguish "valid" from "invalid" responders
- Thus, she would incorrectly conclude that there is a link between the two constructs
Acquiescence Bias: "Nay-Saying"

- We have focused on acquiescence or "yea-saying" bias—where an individual agrees with whatever is presented.
- However, it can also take the form on nonacquiescence bias—"nay-saying", where an individual disagrees with whatever is presented.
- A nay-saying bias can have similar consequences to a yea-saying bias.
- Respondents who obtain low scores on one test will obtain low scores on another test, causing correlations that are artificially higher than they should be.
Acquiescence Bias: Factors that Increase Acquiescence

- Acquiescence bias is a threat to the psychometric quality of a test
- Factors that increase the acquiescence bias include
  1. ambiguous items
  2. long items
  3. large number of items
- Additionally, some people are naturally more susceptible to the bias than others
The response biases that will be discussed in this lecture are:

- Acquiescence Bias
- Extreme and Moderate Responding
- Social Desirability
- Malingering
- Random Responding
- Guessing
Types of Response Biases

• The response biases that will be discussed in this lecture are:
  • Acquiescence Bias
  • Extreme and Moderate Responding
  • Social Desirability
  • Malingering
  • Random Responding
  • Guessing
Extreme and Moderate Responding

- The problem of **extreme** and **moderate response biases** refers to differences in the tendency to use or avoid extreme response options.
- Some people are simply more likely to endorse the extreme of a response scale (even though they do not possess the attribute to that extreme degree)—known as "extreme responding".
- Conversely, some people will choose a response somewhere in the middle to avoid making a strong claim.
- This is known as "moderate responding".
Extreme and Moderate Responding

- Suppose a researcher is interested in the association between emotional distress and spirituality.
  - Emotional distress is assessed using the State-Trait Anxiety Inventory:
    - "I lack self-confidence"
    - 4-point response scale: *almost never, sometimes, often, almost always*
  - Spirituality is assessed using the Spirituality/Religiousness scale:
    - "I am a spiritual person"
    - 5-point response scale: *very inaccurate, moderately inaccurate, neither inaccurate nor accurate, moderately accurate, very accurate*
## Extreme and Moderate Responding

### Types of Response Biases
- Acquiescence Bias
- Extreme and Moderate Responding
- Social Desirability
- Malingering
- Random Responding
- Guessing

### Methods For Coping With Response Bias
- Managing Testing Context
- Managing Test Content
- Specialised Tests

### Table: STAI Items and S/R Items

<table>
<thead>
<tr>
<th>Participant</th>
<th>Bias</th>
<th>True Anxiety</th>
<th>STAI Items</th>
<th>True Spirituality</th>
<th>S/R Items</th>
<th>Total S/R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Extreme</td>
<td>14</td>
<td>4 4 4 4 4 4</td>
<td>16</td>
<td>11</td>
<td>4 4 4 4 4 4 16</td>
</tr>
<tr>
<td>2</td>
<td>Moderate</td>
<td>14</td>
<td>3 3 3 3 3 3</td>
<td>12</td>
<td>12</td>
<td>2 2 2 2 2 2 8</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>12</td>
<td>3 4 3 2 3 2</td>
<td>12</td>
<td>4</td>
<td>1 1 1 1 1 1 4</td>
</tr>
<tr>
<td>4</td>
<td>Moderate</td>
<td>6</td>
<td>2 2 2 2 2 2</td>
<td>8</td>
<td>9</td>
<td>2 2 3 3 3 10</td>
</tr>
<tr>
<td>5</td>
<td>Extreme</td>
<td>6</td>
<td>1 1 1 1 1 1</td>
<td>4</td>
<td>8</td>
<td>1 1 1 1 1 4</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>7</td>
<td>1 3 2 1 1 7</td>
<td>15</td>
<td></td>
<td>3 5 4 3 15</td>
</tr>
</tbody>
</table>

**NOTE:** STAI = State-Trait Anxiety Inventory; S/R = Spirituality/Religiousness scale.
Extreme and Moderate Responding

- Extremity bias can generate artificial differences amongst respondents’ test scores.
- For example, participants 1 and 2 have the same "True Anxiety" levels (this is another "all-seeing" example) but participant 1 has a higher observed anxiety score.
- From our all-seeing vantage point, we can see this is because participant 1 adopted an "extreme" extremity bias.
# Extreme and Moderate Responding

<table>
<thead>
<tr>
<th>Participant</th>
<th>Bias</th>
<th>True Anxiety</th>
<th>STAI Items</th>
<th>Total STAI</th>
<th>True Spirituality</th>
<th>S/R Items</th>
<th>Total S/R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Extreme</td>
<td></td>
<td>4 4 4 4 4</td>
<td>16</td>
<td>11</td>
<td>4 4 4 4 4</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>Moderate</td>
<td></td>
<td>3 3 3 3 3</td>
<td>12</td>
<td>12</td>
<td>2 2 2 2 2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td></td>
<td>3 4 3 2</td>
<td>12</td>
<td>4</td>
<td>1 1 1 1 1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Moderate</td>
<td></td>
<td>2 2 2 2</td>
<td>8</td>
<td>9</td>
<td>2 2 3 3 3</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Extreme</td>
<td></td>
<td>1 1 1 1</td>
<td>4</td>
<td>8</td>
<td>1 1 1 1 1</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td></td>
<td>1 3 2 1</td>
<td>7</td>
<td>15</td>
<td>3 5 4 3 3</td>
<td>15</td>
</tr>
</tbody>
</table>

**NOTE:** STAI = State-Trait Anxiety Inventory; S/R = Spirituality/Religiousness scale.
Extreme and Moderate Responding

- Extremity bias can also obscure true differences amongst respondents’ construct levels
- Consider this time participants 2 and 3
- These participants have different "True Anxiety" levels but identical observed anxiety scores
- This is because participant 2 is reluctant to use the more extreme response option required to warrant her true anxiety level
# Extreme and Moderate Responding

<table>
<thead>
<tr>
<th>Participant</th>
<th>Bias</th>
<th>True Anxiety</th>
<th>STAI Items</th>
<th>Total STAI</th>
<th>True Spirituality</th>
<th>S/R Items</th>
<th>Total S/R</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Extreme</td>
<td>14</td>
<td>4 4 4 4 4</td>
<td>16</td>
<td>11</td>
<td>4 4 4 4 4</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>Moderate</td>
<td>14</td>
<td>3 3 3 3 3</td>
<td>12</td>
<td>12</td>
<td>2 2 2 2 2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>12</td>
<td>3 4 3 2 2</td>
<td>12</td>
<td>4</td>
<td>1 1 1 1 1</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Moderate</td>
<td>6</td>
<td>2 2 2 2 2</td>
<td>8</td>
<td>9</td>
<td>2 2 3 3 3</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Extreme</td>
<td>6</td>
<td>1 1 1 1 1</td>
<td>4</td>
<td>8</td>
<td>1 1 1 1 1</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>7</td>
<td>1 3 2 1 1</td>
<td>7</td>
<td>15</td>
<td>3 5 4 3 3</td>
<td>15</td>
</tr>
</tbody>
</table>

**NOTE:** STAI = State-Trait Anxiety Inventory; S/R = Spirituality/Religiousness scale.
Extreme and Moderate Responding

- These types of problems can produce results that lead to inaccurate research conclusions.
- For example, from our all-knowing vantage point we know that the correlation between "True Anxiety" levels and "True Spirituality" levels is $r = -.04$.
- However, based on the data (Total STAI vs. Total S/R) this correlation is $r = .36$.
- Thus, the extremity problem can foster incorrect research conclusions.
Extreme and Moderate Responding

- These types of problems can produce results that lead to inaccurate research conclusions
- For example, from our all-knowing vantage point we know that the correlation between "True Anxiety" levels and "True Spirituality" levels is $r = -0.04$
- However, based on the data (Total STAI vs. Total S/R) this correlation is $r = 0.36$
- Thus, the extremity problem can foster incorrect research conclusions
Extreme and Moderate Responding

- These types of problems can produce results that lead to inaccurate research conclusions.

- For example, from our all-knowing vantage point we know that the correlation between "True Anxiety" levels and "True Spirituality" levels is $r = -0.04$.

- However, based on the data (Total STAI vs. Total S/R) this correlation is $r = 0.36$.

- Thus, the extremity problem can foster incorrect research conclusions.
Extreme and Moderate Responding

- Bear in mind that the use of an extreme response option is not in itself a problem.

- We want test takers to use response options that reflect their construct levels—this sometimes requires extreme responding.

- It becomes a problem when:
  1. people with identical construct levels differ in their tendency to use moderate and extreme response options.
  2. when people with different construct levels do not differ in their willingness to use moderate and extreme response options.
Extreme and Moderate Responding

- Bear in mind that the use of an extreme response option is not in itself a problem.
- We want test takers to use response options that reflect their construct levels—which sometimes requires extreme responding.
- It becomes a problem when:
  1. people with identical construct levels differ in their tendency to use moderate and extreme response options.
  2. when people with different construct levels do not differ in their willingness to use moderate and extreme response options.
Extreme and Moderate Responding

- Bear in mind that the use of an extreme response option is not in itself a problem.
- We want test takers to use response options that reflect their construct levels—this sometimes requires extreme responding.
- It becomes a problem when:
  1. people with identical construct levels differ in their tendency to use moderate and extreme response options.
  2. when people with different construct levels do not differ in their willingness to use moderate and extreme response options.
Types of Response Biases

- The response biases that will be discussed in this lecture are:
  - Acquiescence Bias
  - Extreme and Moderate Responding
  - Social Desirability
  - Malingering
  - Random Responding
  - Guessing
Types of Response Biases

- The response biases that will be discussed in this lecture are:
  - Acquiescence Bias
  - Extreme and Moderate Responding
  - Social Desirability
  - Malingering
  - Random Responding
  - Guessing
Social Desirability

- **Social desirability bias** (also known as *faking good*) is the tendency for a person to respond in a way that seems socially appealing, regardless of his or her true characteristics.

- Recall the item, "Have you ever stolen anything from your employer?"

- If you wanted the job, you would be tempted to provide responses that would appeal to the employer.

- That is, you would want to appear to have integrity, high levels of conscientiousness, and honesty.

- If you do respond in such a way—and you do not actually possess these characteristics—you are exhibiting a social desirability response bias.
Social Desirability

- The textbook says that social desirability bias "can diminish the reliability and validity of the measurement process"
- Social desirability bias does indeed compromise the validity of test score interpretations
- However, it is unlikely to affect reliability
- This is because people would be expected to engage in socially desirable responding across all items, consistently
Social Desirability

- Social desirability bias can be affected by at least three sources:
  1. Test content
  2. Test context
  3. Personality of the respondent
Social Desirability

- Social desirability bias can be affected by at least three sources:
  1. Test content
  2. Test context
  3. Personality of the respondent
Social Desirability: Test Content

- Some psychological constructs have greater implications for social appeal than do others.
- For example, personality characteristics such as psychological well being (vs. psychological distress) or honesty (vs. deceitfulness) might be closely linked to social desirability.
- Well-being and honesty are probably viewed as more socially desirable than distress and deceitfulness.
- Other characteristics—such as extraversion (vs. introversion)—might be less affected by social desirability.
Social Desirability

- Social desirability bias can be affected by at least three sources:
  1. Test content
  2. Test context
  3. Personality of the respondent
Social Desirability

- Social desirability bias can be affected by at least three sources:
  1. Test content
  2. Test context
  3. Personality of the respondent
Social Desirability: Test Context

- Socially desirable responding might be more likely to occur in contexts in which respondents can be identified than in contexts in which they are anonymous.
- When respondents can be linked to their responses, they might be more likely to provide responses that are socially appealing.
- Additionally, socially desirable responding may be more likely in testing contexts where important consequences hinge on testing outcomes.
- The hiring example is a case in point.
- Socially desirable responding is less of a problem when there are no important testing consequences (though it is still a problem).
Social desirability bias can be affected by at least three sources:

1. Test content
2. Test context
3. Personality of the respondent
Social Desirability

- Social desirability bias can be affected by at least three sources:
  1. Test content
  2. Test context
  3. Personality of the respondent
Another source of socially desirable responding is the personality of the test taker.

Some people have a personality type that disposes them to engaging in socially desirable responding.

For example, individual differences in the "need for autonomy" are known to correlate negatively with socially desirable responding.

People high in need for autonomy are less sensitive to social disapproval.

Thus, they score high in need for autonomy and low in socially desirable responding.
Social Desirability: Implications

- Social desirability bias has implications for:
  1. applied psychology
  2. behavioural research
Social desirability bias has implications for:

1. **applied psychology**
2. **behavioural research**
Social Desirability: Applied Psychology

- Many personality tests are "transparent" in that it is clear what types of qualities they are measuring (high face validity).
- In these tests, it is easy to "fake" having desirable qualities.
- Such faking behaviour can be quite common, especially in the context of job applications.
- If applicants exaggerate their positive qualities and minimise their negative qualities, then hiring decisions can be compromised.
- When faking takes place, a person hired under a testing based selection process is likely to be a "faker".
Social desirability bias has implications for:

1. applied psychology
2. behavioural research
Social Desirability: Implications

- Social desirability bias has implications for:
  1. applied psychology
  2. behavioural research
Social Desirability: Behavioural Research

- Socially desirable responding can have a similar effect on correlations between scores as was demonstrated for the case of acquiescence.

- The textbook mentions the example of positive affect and relationship quality.

- People who tend to engage in socially desirable responding will report that they have higher levels of positive affect than they in fact do.

- Similarly, they will also rate the quality of their relationships as better than they in fact are.

- This causes the correlation between the two scores to be higher than it really is (see Table 10.3, p.283).
Social Desirability: The Nature of Socially Desirable Responding

- Del Paulhus has done much work exploring social desirability bias as an aspect of personality.
- An implication of his research is that there may be multiple forms of social desirability bias.
- He proposed two processes by which socially desirable responding occurs:
  1. Impression management
  2. Self-deception
Social Desirability: The Nature of Socially Desirable Responding

- Del Paulhus has done much work exploring social desirability bias as an aspect of personality.
- An implication of his research is that there may be multiple forms of social desirability bias.
- He proposed two processes by which socially desirable responding occurs:
  1. Impression management
  2. Self-deception
Social Desirability: Impression Management

- This is a **conscious** process, where test takers intentionally attempt to appear socially desirable.
- For example, a job applicant might exaggerate their desirable qualities and minimise their undesirable qualities.
- Impression management is thought to be a psychological **state** that depends on situational demands.
- Thus, conscious efforts to appear socially desirable usually arise in particular contexts (e.g., completing a personality inventory for a job interview).
Social Desirability: The Nature of Socially Desirable Responding

- Del Paulhus has done much work exploring social desirability bias as an aspect of personality.
- An implication of his research is that there may be multiple forms of social desirability bias.
- He proposed two processes by which socially desirable responding occurs:
  1. Impression management
  2. Self-deception
Social Desirability: The Nature of Socially Desirable Responding

- Del Paulhus has done much work exploring social desirability bias as an aspect of personality.
- An implication of his research is that there may be multiple forms of social desirability bias.
- He proposed two processes by which socially desirable responding occurs:
  1. Impression management
  2. Self-deception
Social Desirability: Self-Deception

- This is thought to be an unconscious process
- People who engage in self-deception tend to actually believe their overestimation of their psychological characteristics
- It is correlated positively with narcissism
- Unlike impression management, self-deception is thought to reflect a psychological trait that is stable across time and context
- Thus, some people are naturally more predisposed toward self-deception than others
Types of Response Biases

- The response biases that will be discussed in this lecture are:
  - Acquiescence Bias
  - Extreme and Moderate Responding
  - Social Desirability
  - Malingering
  - Random Responding
  - Guessing
Types of Response Biases

- The response biases that will be discussed in this lecture are:
  - Acquiescence Bias
  - Extreme and Moderate Responding
  - Social Desirability
  - Malingering
  - Random Responding
  - Guessing
Malingering

- This is the opposite problem to social desirability
- Respondents might sometimes try to exaggerate their psychological problems
- For example, in some applied testing contexts, test takers may be motivated to appear more cognitively impaired or psychologically disturbed than they truly are
- This is known as **malingering** or **faking bad**

  Malingering is particularly problematic in testing situations where individuals stand to benefit from their faking behaviour
  - This includes criminal competency hearings, worker compensation claims, and personal injury examinations
Malingering

• Malingering compromises the quality of psychological assessment

• For example, suppose an individual involved in a car accident takes a neuropsychological assessment that includes measures of his attention and memory

• If the individual deliberately tries to perform poorly on these measures—to reap financial benefits—then the accuracy and fairness of the diagnosis and decisions made on the basis of those measures will be compromised

• Failure to consider the impact of malingering potentially carries high costs for insurers, disability systems, and society at large
Malingering is common in applied testing situations.

It occurs in between 7% to 27% of general psychological evaluations and between 31% to 45% of forensic evaluations (e.g., criminal competence, disability hearings etc.)

There is in fact evidence that lawyers coach clients in methods to detect malingering.

Malingering is therefore a legitimate concern in psychological testing, especially in situations where test takers stand to reap benefits by exaggerating their psychological problems.
Types of Response Biases

- The response biases that will be discussed in this lecture are:
  - Acquiescence Bias
  - Extreme and Moderate Responding
  - Social Desirability
  - Malingering
  - Random Responding
  - Guessing
Types of Response Biases

- The response biases that will be discussed in this lecture are:
  - Acquiescence Bias
  - Extreme and Moderate Responding
  - Social Desirability
  - Malingering
  - Random Responding
  - Guessing
Careless or Random Responding

- Whether due to carelessness or lack of motivation, sometime test takers respond randomly or semi-randomly
- For example, an individual taking a test in which the items are answered on a 5-point scale (1 = strongly disagree, 5 = strongly agree) might cycle through the response scale:
  - That is, select strongly disagree for Item 1, disagree for Item 2, neutral for Item 3, agree for Item 4, strongly agree for Item 5, and then wrap back around for Item 6
- This produces scores that are meaningless with regards the construct being measured
The response biases that will be discussed in this lecture are:

- Acquiescence Bias
- Extreme and Moderate Responding
- Social Desirability
- Malingering
- Random Responding
- Guessing
Types of Response Biases

• The response biases that will be discussed in this lecture are:
  • Acquiescence Bias
  • Extreme and Moderate Responding
  • Social Desirability
  • Malingering
  • Random Responding
  • Guessing
Guessing

- Aptitude and achievement tests are designed so that specific responses are correct and others incorrect.
- Scores on these tests are often consequential in that they are used for some form of selection (e.g., college admission).
- Respondents may be motivated to guess on such tests, especially if they are based on a limited set of options (e.g., multiple-choice questions).
Guessing

- Guessing can compromise the quality of test scores
- A correct guess artificially inflates a test taker’s score relative to his or her true score
- Guessing therefore produces test scores that are inconsistent with the true score differences, which has an adverse effect on test reliability
Methods For Coping With Response Bias

- There are at least three general kinds of **strategies**:
  1. Manage the testing context
  2. Manage the test content and/or scoring
  3. Use specially designed "bias" tests

- Additionally, there at least three **goals** that these strategies are intended to accomplish:
  1. Minimize the existence of response biases
  2. Minimize the effects of responses biases
  3. Detect bias and intervene in some way
Methods For Coping With Response Bias

• There are at least three general kinds of strategies:
  1. Manage the testing context
  2. Manage the test content and/or scoring
  3. Use specially designed "bias" tests

• Additionally, there are at least three goals that these strategies are intended to accomplish:
  1. Minimize the existence of response biases
  2. Minimize the effects of responses biases
  3. Detect bias and intervene in some way
Minimising the Existence of Bias by Managing Testing Context

- Perhaps the best way to cope with response biases is to prevent them occurring in the first place.
- The occurrence of response biases can be minimised by managing the testing context.
- For example, respondents may be less likely to engage in socially desirable responding, if they are convinced that they are responding **anonymously**.
- The rationale is that anonymity allows respondents to feel comfortable in honestly admitting to undesirable attributes (about themselves or others).
- The downside is that although anonymity increases honest responding, it may also increase random responding.
Minimising the Existence of Bias by Managing Testing Context

- When people are tired, they may increase biased responding.
- Another strategy therefore is to create a testing situation that minimises respondent fatigue, stress, distraction, or frustration.
- A final strategy is known as the "bogus pipeline technique"—inform participants that false responding can be detected by the test that is being administered.
- The mere possibility of detection can convince some respondents to be more honest and unbiased.
- This strategy is a potential solution to malingering.
Methods For Coping With Response Bias

• There are at least three general kinds of strategies:

1. Manage the testing context
2. Manage the test content and/or scoring
3. Use specially designed "bias" tests

• Additionally, there at least three goals that these strategies are intended to accomplish:

1. Minimize the existence of response biases
2. Minimize the effects of responses biases
3. Detect bias and intervene in some way
Methods For Coping With Response Bias

- There are at least three general kinds of strategies:
  1. Manage the testing context
  2. Manage the test content and/or scoring
  3. Use specially designed "bias" tests

- Additionally, there are at least three goals that these strategies are intended to accomplish:
  1. Minimize the existence of response biases
  2. Minimize the effects of responses biases
  3. Detect bias and intervene in some way
Minimising the Existence of Bias by Managing Test Content

- Test developers sometimes choose specific kinds of items or specific kinds of response formats to help minimize some biases:
  1. Write clear, concise, and unambiguous items
     - this reduces frustration and distraction that might cause low motivation and biased responding
  2. Write items that are neutral in terms of social desirability:
     - instead of "I am a hostile person"
     - use "I am sometimes less friendly than other people"
  3. Forced choice-formats
Minimising the Existence of Bias by Managing Test Content

• Test developers sometimes choose specific kinds of items or specific kinds of response formats to help minimize some biases:
  1. Write clear, concise, and unambiguous items
     • this reduces frustration and distraction that might cause low motivation and biased responding
  2. Write items that are neutral in terms of social desirability:
     • instead of "I am a hostile person"
     • use "I am sometimes less friendly than other people"
  3. Forced choice-formats
Minimising the Existence of Bias by Managing Test Content

- Test developers sometimes choose specific kinds of items or specific kinds of response formats to help minimize some biases:
  1. Write clear, concise, and unambiguous items
     - this reduces frustration and distraction that might cause low motivation and biased responding
  2. Write items that are neutral in terms of social desirability:
     - instead of "I am a hostile person"
     - use "I am sometimes less friendly than other people"
  3. Forced choice-formats

Test developers sometimes choose specific kinds of items or specific kinds of response formats to help minimize some biases:

1. Write clear, concise, and unambiguous items
   - this reduces frustration and distraction that might cause low motivation and biased responding
2. Write items that are neutral in terms of social desirability:
   - instead of "I am a hostile person"
   - use "I am sometimes less friendly than other people"
3. Forced choice-formats
Minimising the Existence of Bias by Managing Test Content

- Test developers sometimes choose specific kinds of items or specific kinds of response formats to help minimize some biases:
  1. Write clear, concise, and unambiguous items
     - this reduces frustration and distraction that might cause low motivation and biased responding
  2. Write items that are neutral in terms of social desirability:
     - instead of "I am a hostile person"
     - use "I am sometimes less friendly than other people"
  3. Forced choice-formats
Minimising the Existence of Bias by Managing Test Content

- Test developers have used "forced-choice" formats to minimise the existence of social desirability bias.
- Forced-choice items are items that present two characteristics and require that respondents endorse one and only one of them.
- For example, an item on a personality test might present the two equally socially desirable characteristics "friendly" and "assertive" and test takers choose the one that best describes them personally.
- Another item might include characteristics that are equally undesirable (e.g., "timid" and "argumentative").
- The forced-choice format prevents test takers from simply picking the more desirable choice.
Minimising the Existence of Bias by Managing Test Content or Scoring

- Test developers can also design tests to minimise the extent of the extreme responding problem.
- For example, they can provide only two responses for each item.
- On a personality test, a test taker might be asked to respond "Yes" or "No" to whether they consider themselves "friendly".
- Such a format eliminates extremity bias by removing any "extreme" response options altogether.
Methods For Coping With Response Bias

There are at least three general kinds of strategies:

1. Manage the testing context
2. Manage the test content and/or scoring
3. Use specially designed "bias" tests

Additionally, there are at least three goals that these strategies are intended to accomplish:

1. Minimize the existence of response biases
2. Minimize the effects of responses biases
3. Detect bias and intervene in some way
Methods For Coping With Response Bias

- There are at least three general kinds of *strategies*:
  1. Manage the testing context
  2. Manage the test content and/or scoring
  3. Use specially designed "bias" tests

- Additionally, there at least three *goals* that these strategies are intended to accomplish:
  1. Minimize the existence of response biases
  2. Minimize the effects of responses biases
  3. Detect bias and intervene in some way
Minimising the Effects of Bias by Managing Test Content or Scoring

- In reality, it is very difficult to **eliminate** the existence of response bias
- Instead, test developers typically try to **minimise** the effects of response bias
- A common method used to help minimise the effects of acquiescence bias is to create *balanced scales*
- A balanced scale consists of an equal number of positively keyed and negatively keyed items
Recall the job satisfaction scale presented in the discussion of acquiescence bias:

1. "I really enjoy my work"
2. "I find my work personally fulfilling"
3. "In general, I am satisfied with the day-to-day aspects of my job"
4. "There is very little that I would change about my job"

All four items are positively keyed, which might encourage acquiescence bias.
Minimising the Effects of Bias by Managing Test Content or Scoring

- To minimise the potential effects of acquiescence bias, we might use the following revised set of items:
  1. "I really enjoy my work"
  2. "I do not find my work personally fulfilling"
  3. "In general, I am satisfied with the day-to-day aspects of my job"
  4. "There is much I would change about my job"

- The revised scale is balanced—it contains equal numbers of positively keyed and negatively keyed statements.

- Balanced scales allow a test user to identify an acquiescent responder.

- Such a responder would agree with all four items, despite the fact that some agreements indicate a high level of satisfaction, but others indicate a low level.
Minimising the Effects of Bias by Managing Test Content or Scoring

- Note that there are some practicalities when it comes to analysing balanced scales
- Before you can do any analyses, you must "reverse-score" each of the negatively keyed items
- That is, you want high scores to be indicative of the attribute of interest
- In this example, someone who responded a "strongly disagree" to Item 2 would get a score of 7 for that item (rather than 1)
- The same goes for item 4
Minimising the Effects of Bias by Managing Test Content or Scoring

- Test users can also use specialised scoring procedures to minimise guessing
- Achievement tests like the SAT adjust test scores to account for guessing
- Specifically, in the SAT an item answered incorrectly results in a subtraction of .25 of a point
- An item left unanswered results in no such points deduction
- This differential scoring procedure discourages guessing behaviour amongst SAT test takers
There are at least three general kinds of strategies:

1. Manage the testing context
2. Manage the test content and/or scoring
3. Use specially designed "bias" tests

Additionally, there are at least three goals that these strategies are intended to accomplish:

1. Minimize the existence of response biases
2. Minimize the effects of responses biases
3. Detect bias and intervene in some way
Methods For Coping With Response Bias

There are at least three general kinds of strategies:

1. Manage the testing context
2. Manage the test content and/or scoring
3. Use specially designed "bias" tests

Additionally, there are at least three goals that these strategies are intended to accomplish:

1. Minimize the existence of response biases
2. Minimize the effects of responses biases
3. Detect bias and intervene in some way
Managing Test Content to Detect Bias and Intervene

- Some inventories incorporate so-called *validity scales*.
- Validity scales are designed to specifically measure the degree to which people may be engaging in a particular type of response bias.
- Perhaps the most famous self-report inventory is the Millon Clinical Multiaxial Inventory (MMPI).
- The MMPI is used by psychiatrists and psychologists to measure personality (but more along the psychopathological ends of the spectrum).
- To help determine whether the respondent’s scores are interpretable, they consult the validity scales.
Managing Test Content to Detect Bias and Intervene

- The two main validity scales within the MMPI are the L scale and the F scale

- The L(ie) scale:
  - consists of 15-items that describe "minor flaws and weaknesses to which most people are willing to admit"
  - a high score represents denial of these common characteristics which is indicative of social desirability bias

- The F (Infrequency) scale:
  - consists of 64-items that are endorsed by very few respondents
  - a high score represents some form of deviant responding—random responding, malingering, acquiescence, or serious psychopathology
The MMPI gives instructions on what to do if someone has elevated scores on the validity scales:

- In some cases, you might totally exclude any interpretation of the scores.
- In others, you might interpret them cautiously.
- Alternatively, you might use statistical procedures to "correct" the original scores into more valid scores.
Methods For Coping With Response Bias

• There are at least three general kinds of strategies:
  1. Manage the testing context
  2. Manage the test content and/or scoring
  3. Use specially designed "bias" tests

• Additionally, there are at least three goals that these strategies are intended to accomplish:
  1. Minimize the existence of response biases
  2. Minimize the effects of responses biases
  3. Detect bias and intervene in some way
There are at least three general kinds of strategies:

1. Manage the testing context
2. Manage the test content and/or scoring
3. Use specially designed "bias" tests

Additionally, there are at least three goals that these strategies are intended to accomplish:

1. Minimize the existence of response biases
2. Minimize the effects of responses biases
3. Detect bias and intervene in some way
Using Specialised Tests to Detect Bias and Intervene

- Unlike the validity scales found in a test such as the MMPI, there are also tests dedicated to the measurement of socially desirable responding.

- Two of the most famous tests designed for this purpose are the Marlowe-Crowne Social Desirability Scale and the Balanced Inventory for Desirable Responding (see the textbook for a description p.297).

- These scales exist independently of any other scale, so they can be used in a variety of contexts (applied or research).

- They can be used in conjunction with other tests to establish the degree to which social desirability bias may have affected scores on those tests.
Using Specialised Tests to Detect Bias and Intervene

- The Dot Counting Test is an example of a test specifically designed to detect malingered cognitive impairment.
- Consists of 12 cards with half of them containing dots presented in a grouped fashion and the other half presented in a random fashion.
- Malingering is suspected when it takes people just as long to count the grouped dots as the random dots.
- Note that the cognitive competencies required to perform this task are so primitive they should be present in all but the most severely brain injured patients.
- This is why it constitutes a measure of malingering.
Using Specialised Tests to Detect Bias and Intervene

• The Dot Counting Test is an example of test specifically designed to detect malingered cognitive impairment.

• Consists of 12 cards with half of them containing dots presented in a grouped fashion and the other half presented in a random fashion.

• Malingering is suspected when it takes people just as long to count the grouped dots as the random dots.

• Note that the cognitive competencies required to perform this task are so primitive they should be present in all but the most severely brain injured patients.

• This is why it constitutes a measure of malingering.
Using Specialised Tests to Detect Bias and Intervene

Types of Response Biases
- Acquiescence Bias
- Extreme and Moderate Responding
- Social Desirability
- Malingering
- Random Responding
- Guessing

Methods For Coping With Response Bias
- Managing Testing Context
- Managing Test Content
- Specialised Tests

Tese types of cards should take less time to count.
Using Specialised Tests to Detect Bias and Intervene

- The Dot Counting Test is an example of test specifically designed to detect malingered cognitive impairment
- Consists of 12 cards with half of them containing dots presented in a grouped fashion and the other half presented in a random fashion
- Malingering is suspected when it takes people just as long to count the grouped dots as the random dots
- Note that the cognitive competencies required to perform this task are so primitive they should be present in all but the most severely brain injured patients
- This is why it constitutes a measure of malingering
Terminology: Response Sets vs. Response Styles

- **Response Sets:**
  - Are temporary and due to factors such as circumstances surrounding the testing or the test itself

- **Response Styles:**
  - Are considered relatively stable and enduring traits that are observed across tests and testing situations
  - Some people are more concerned about appearing socially desirable than others

- Response biases can arise from response sets and response styles