

## Issues to consider when drafting questions

### Social Desirability Bias

Wanting to make a good impression

- "How often do you drive a car after drinking alcoholic beverages? Frequently,
   Occasionally, Seldom, Never, or Don't Know." (Dillman & Tarnai, 1991)
  - interviewer-administered telephone survey 63% said "never"
  - > self-administered paper survey 52% responded "never"
- "How would you describe your current health? Excellent, Good, Fair, or Poor."
  - fewer respondents choose "excellent" in self-administered surveys compared to interviews (Biemer, 1997; Hochstim, 1967).
    - conventional American greeting "How are you?" -> always "Fine"

### Social Desirability Bias

- Kreuter et al (2008):
  - ▶ 20% of web survey respondents denied having ever received a D or F in college when in fact they had received one of these grades.

- Catania et al (1996):
  - both men and women were more likely to report engaging in extramarital sex when interviewed by a same sex interviewer than when interviewed by an opposite sex interviewer.

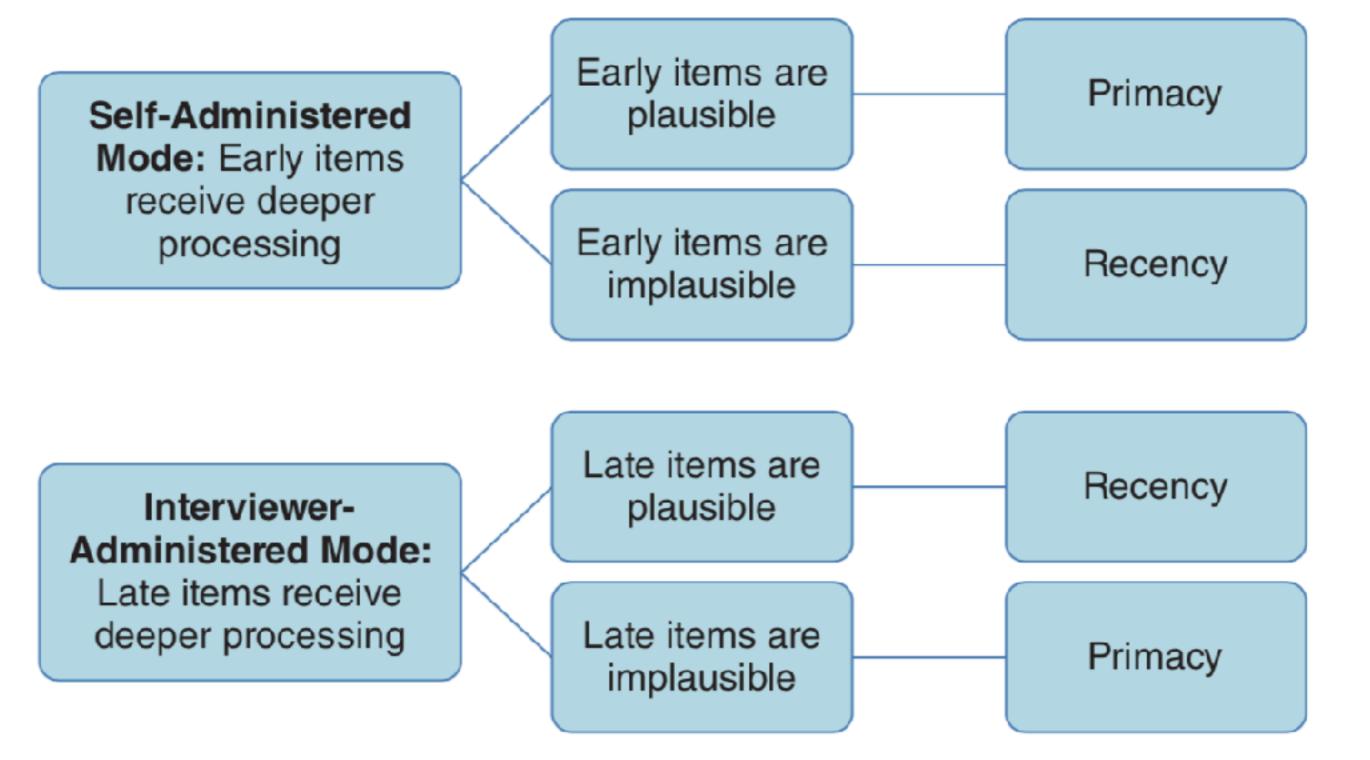
### Acquiescence

The tendency to agree with someone rather than disagree

- Schuman and Presser (1981):
  - 60% agreed "Individuals are more to blame than social conditions for crime and lawlessness in this country."
  - > 57% of a control group agreed with the exact reverse "Social conditions are more to blame than individuals for crime and lawlessness in this country."

### Primacy and Recency

FIGURE 4.4 Cognitive elaboration model of response order effects.



Source: Adapted from "A Cognitive Model of Response Order Effects in Survey Measurement," by N. Schwarz, H. J. Hippler, and E. Noelle-Neumann, 1992, in N. Schwarz and S. Sudman (Eds.), Context Effects in Social and Psychological Research (pp. 187–201), New York, NY: Springer-Verlag.

Tendency to more frequently choose from among the first / last categories offered regardless of their content

### Anchoring

Occurs when an early response option forms a standard of comparison for later response options.

- Noelle-Neumann (1970): "Which food is more typically German?"
  - "potatoes" and "rice" 30% said "potatoes"
  - "rice" and "potatoes" 48% said "potatoes"





## The anatomy of a survey question

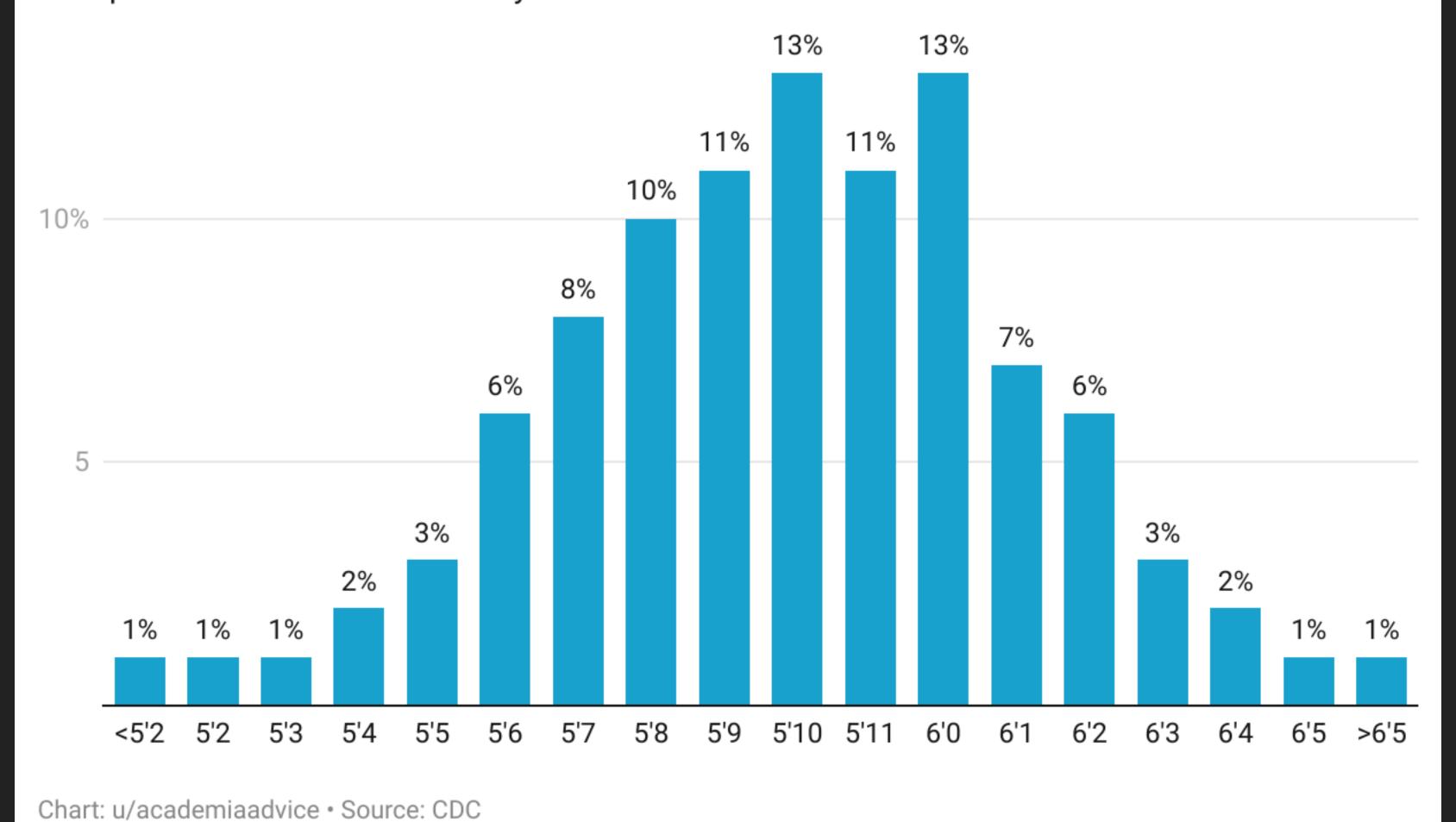
## Types of Survey Questions

Open-ended questions					
Question stem	What is the most important problem facing Nebraska today?				
Answer space					
Question stem with verbal and numeric instructions	How many years have you lived in Nebraska? Please report only whole numbers. For example, if you have lived in Nebraska 20 months, please round to 2 years.				
Answer space with verbal and symbolic instruction	Years lived in Nebraska				
Closed-ended ordinal question					
Question stem	Overall, how satisfied are you with living in Nebraska?				
Answer choices	<ul> <li>Completely satisfied</li> <li>Very satisfied</li> <li>Somewhat satisfied</li> <li>Not too satisfied</li> <li>Not at all satisfied</li> </ul>				

## Types of Survey Questions

### **Body Height Reported by U.S. Men**

As part of a comprehensive health survey, the U.S. CDC asked roughly 200,000 adult men in 2021 this question: "About how tall are you without shoes?"

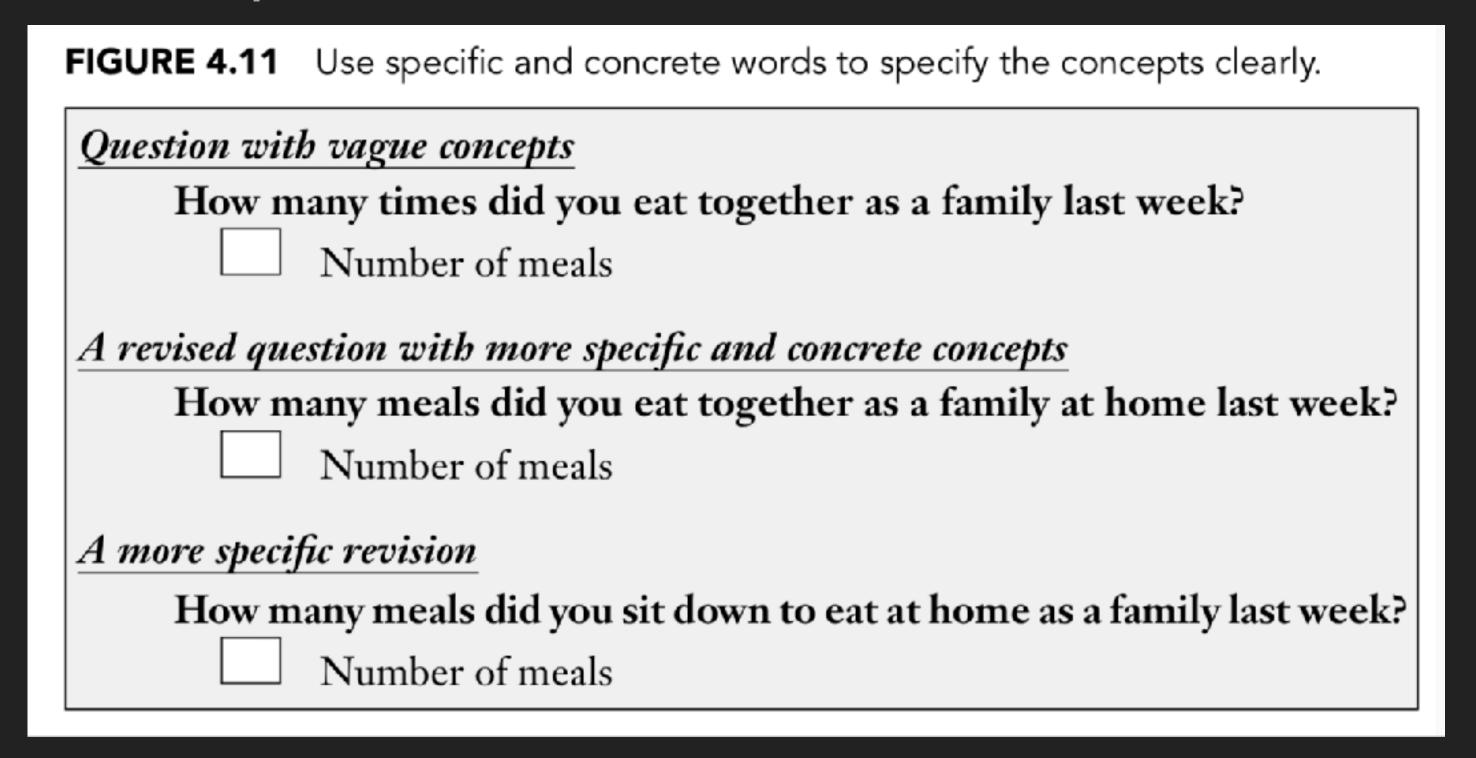


## Types of Survey Questions

Closed-ended nominal question						
Question stem	What is your current marital status?					
Answer choices	<ul> <li>Married</li> <li>Living with a partner</li> <li>Divorced</li> <li>Separated</li> <li>Widowed</li> <li>Never married</li> </ul>					
Partially closed-ended question						
Question stem	What are your favorite women's sports at the University of Nebraska?					
Answer choices	<ul> <li>□ Basketball</li> <li>□ Cross Country</li> <li>□ Gymnastics</li> <li>□ Soccer</li> <li>□ Softball</li> <li>□ Swimming and Diving</li> <li>□ Tennis</li> <li>□ Volleyball</li> <li>□ Other: Please specify</li> </ul>					

### Use Specific and Concrete Words To Specify the Concepts Clearly

- Questions may be factual, but interesting ones usually involve concepts
- What are the concepts?



Use previously-validated scales where possible

### Guidelines for Forming Questions

- Choose the appropriate question format
- Make sure the question applies to the respondent
- Ask one question at a time
- Make sure the question is technically accurate
- Use simple and familiar words
- Use specific and concrete words to specify the concepts clearly
- Use as few words as possible to pose the question
- Use complete sentences that take a question form, and use simple sentence structures
- Make sure "yes" means yes and "no" means no
- Organize questions to make it easier for respondents to comprehend the response task

## How to Write Open- and Closed-Ended Questions

### Open Ended Questions — Wording Specificity

**FIGURE 5.1** How specificity of question wording affects reports about when students began their studies.

	% Reporting	% Reporting
Question Wording (Telephone)	<b>Month and Year</b>	Season/Semester
When did you begin your studies at Washington State University?	13.4	57.3
What date did you begin your studies at Washington State University?	49.5	32.3
What month and year did you begin your studies at Washington State University?	83.7	11.0

Source: "Helping Respondents Get It Right the First Time: The Influence of Words, Symbols, and Graphics in Web Surveys," by L. M. Christian, D. A. Dillman, and J. D. Smyth, 2007b, *Public Opinion Quarterly*, 71(1), pp. 113–125.

### Open Ended Questions — Provide Extra Motivation To Respond

In your own words, how would you describe your adviser(s)?"

Vs

"This question is very important to understanding the Washington State University student experience. Please take your time answering it.

In your own words, how would you describe your adviser(s)?"

### Open Ended Questions — Use Nondirective Probes To Obtain More Information

- Smyth et al. (2007b) "What businesses would you most like to see in the Pullman and Moscow area that are currently not available?"
  - A random half of students received a follow-up probe asking, "Are there any others?"
- No probe: average of 1.8 businesses
- Probe: average of 2.4 businesses

# Open Ended Questions — the Type of Probe Used Will Strongly Impact the Amount and Type of Information Received

Smyth et al. (2006) – "In your own words, how would you describe your adviser or advisers?"

### Probes:

- "Is there anything else?" 18% responded; most said "no"
- "Can you tell me more about that?" 82% responded with additional information, including new ideas or themes as well as elaboration on previously reported themes

### Closed Ended Questions — Acquiescence

#### **Poor Designs**

Do you favor congressional term limits of four years?

- O Favor
- O Oppose

How satisfied are you with the overall service you have received from your financial consultant?

- O Very satisfied
- O Somewhat satisfied
- Somewhat dissatisfied
- O Very dissatisfied

### Closed Ended Questions — Acquiescence

State both positive and negative sides in the question stem. Improved Designs Poor Designs Do you favor congressional term Do you favor or oppose congressional limits of four years? term limits of four years? O Favor O Favor O Oppose O Oppose How satisfied are you with the How satisfied or dissatisfied are you with the overall service you have overall service you have received from your financial consultant? received from your financial consultant? O Very satisfied Somewhat satisfied O Very satisfied O Somewhat satisfied Somewhat dissatisfied Very dissatisfied Somewhat dissatisfied O Very dissatisfied

### Closed Ended Questions — Primacy

### FIGURE 5.7 Subtraction effects in multiple-answer questions.

Which of the following resources have you used at WSU? Please check all that apply. %				
Original Order Endors	ing Reverse Order Endorsing			
Libraries 95	Counseling Services			
Computer Labs	Library Instruction 52			
Student Health Center	Campus-Sponsored Tutoring			
Academic Advising	Career Services			
Student Recreation Center	Internet/E-Mail Access			
Internet/E-Mail Access	Student Recreation Center			
Career Services	Academic Advising			
Campus-Sponsored Tutoring	Student Health Center			
Library Instruction 20	Computer Labs			
Counseling Services	Libraries93			

### Closed Ended Questions — Forced Choice

Check-all-that-apply formatted question		
Which of the following items do you have? Please check all that apply.		
☐ Desktop computer		
☐ Laptop computer		
□ Cell phone		
□ E-reader		
☐ Tablet computer		
□ iPod or MP3 player		

### Closed Ended Questions — Forced Choice

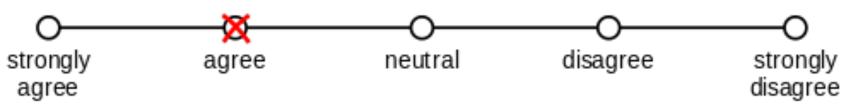
FIGURE	5.12	2 Replacing check-all-that-apply questions with a forced-choice format.					
Check-all-that-apply formatted question							
Which of the following items do you have? Please check all that apply.							
☐ Desktop computer ☐ Laptop computer ☐ Cell phone ☐ E-reader ☐ Tablet computer ☐ iPod or MP3 player  A revision converting to the forced-choice format  Do you have each of the following items or not?							
	Yes						
	O	No O	Desktop computer				
	O	Ö	Laptop computer				
	0	O	Cell phone				
	O	O	E-reader				
	O	O	Tablet computer				
	$\circ$	$\circ$	iPod or MP3 player				

## Understanding Likert scales better

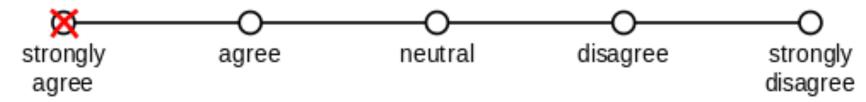
## Likert Scale (Pronounced Lick-Ert)

### Website User Survey

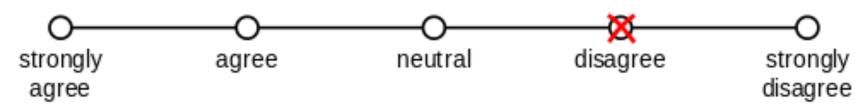
1. The website has a user friendly interface.



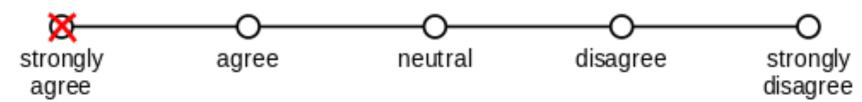
2. The website is easy to navigate.



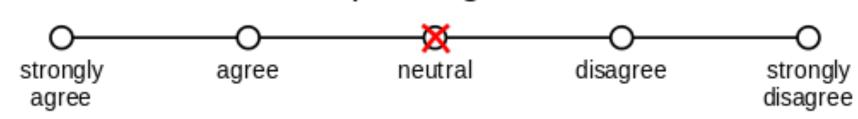
3. The website's pages generally have good images.



4. The website allows users to upload pictures easily.



5. The website has a pleasing color scheme.



"Example Likert Scale using five Likert Items pertaining to Wikipedia" CC-BY-SA-3.0 Wikipedia

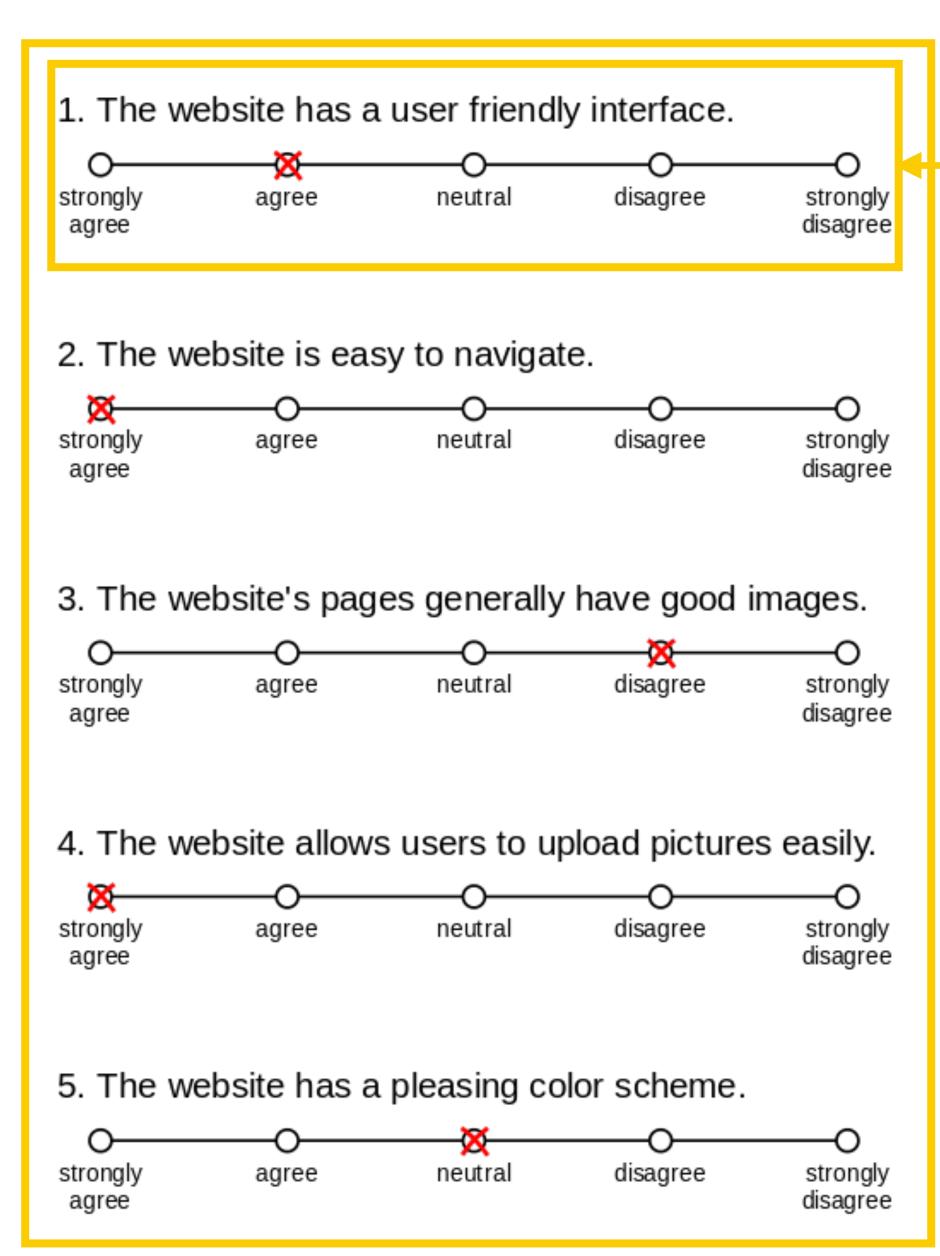
# Likert Scale (Pronounced Lick-Ert)

"strongly disagree" = 1

• • •

"strongly agree" = 5

### Website User Survey



Likert item

Likert Scale

"Example Likert Scale using five Likert Items pertaining to Wikipedia" CC-BY-SA-3.0 Wikipedia

### Concerns: Reliability - Is Whatever Is Being Measured Reliably Found?

- Statistical or internal reliability (aka "consistency"):
  - The extent to which the items of a scale are consistent with each other.
  - Typically evaluated using Cronbach's a
    - A good scale should have a good split-half correlation.
    - Cronbach's a is effectively the average of all the possible split-half correlations.
- Reliability over time (aka "stability")
  - Over time, people should answer the same items the same way.
    - Especially for stable attributes like attitude or personality

# Concerns: Validity - Does Whatever Is Being Measured Correspond to the Concept That We Say It Does?

- Do scales with different response formats agree with each other?
  - e.g., similar mean values with different items
- Do items correlate with reference criteria of what the concept means?

### Should Items Have a Midpoint?

I like massively multiplayer online role-playing games (MMORPG)
 Strongly
 Disagree
 Neither
 Agree
 Agree

### Should Items Have a Midpoint?

Really ambivalent about their attitude or simply do not have an opinion?

```
I like massively multiplayer online role-playing games (MMORPG)
                                                                     Strongly
        Strongly
                        Disagree
                                         Neither
        Disagree
                                                                      Agree
                                                         Agree
    I like massively multiplayer online role-playing games (MMORPG)
                                                        Strongly
        Strongly
        Disagree
                        Disagree
                                                         Agree
                                         Agree
3.
    I like massively multiplayer online role-playing games (MMORPG)
                                                        Strongly
        Strongly
                                                                        No
                        Disagree
        Disagree
                                                         Agree
                                                                     Opinion
                                         Agree
                 Three formats of Likert item: 1. with midpoint; 2. no midpoint;
                   3. no midpoint but option to have no opinion
```

### One Problem of Having a Midpoint Is That of Acquiescence Bias

- A neutral option means respondents can comfortably avoid disagreeing even if they do actually disagree.
- Garland (1991) experiment:
  - 4-point (no midpoint) vs 5-point items
  - scale scores were higher for the 5-point items
  - But Nadler et al (2015) found no statistically significant differences

### But Bias Could Also Occur in the Opposite Direction

4-point item: Forcing people to respond in a way that they did not truly feel
would introduce an element of randomness into the answers.

### Likert Item Design Summary

- Likert items are pretty robust to variations in response format.
  - Midpoint vs not
  - 5 vs 7 options
    - Strongly disagree, Disagree, Neither, Agree, Strongly agree
    - Strongly disagree, Moderately disagree, Slightly disagree, Neither, Slightly agree,
       Moderately agree, Strongly agree
  - All-labeled vs just end-points labeled
    - Strongly Disagree, Disagree, Neither, Agree, Strongly Agree
    - Strongly Disagree, 1, 2, 3, 4, 5, Strongly Agree
- Use larger scales over single items.

### Example: Questionnaire to Seventh' Graders Pleasure in Writing

- I love writing.
- Writing is my favorite school subject.
- When I write, I feel well.
- I hate writing.
- I write as soon as I get the chance.
- I make sure that I have to write as less as possible.
- I write more than my class mates.
- When I write, I prefer to do something different.
- Writing gives me pleasure.
- I just write, when I can get a good grade for it.
- Writing is boring.

- I like different kinds of writing.
- When I have the opportunity to determine on my own what I do in the Dutch class, I usual do a writing task.
- I write even if the teacher does not assign a writing task.
- I would like to spend more time on writing.
- Writing is a waste of time.
- I always look forward to writing lessons.
- I write because I have to at school.
- I like it to write down my thoughts.
- I would like to write more at school.

### Example: Questionnaire to Seventh' Graders Pleasure in Writing

- I love writing.
- Writing is my favorite school subject.
- When I write, I feel well.
- I hate writing.
- I write as soon as I get the chance.
- I make sure that I have to write as less as possible.
- I write more than my class mates.
- When I write, I prefer to do something different.
- Writing gives me pleasure.
- I just write, when I can get a good grade for it.
- Writing is boring.

- I like different kinds of writing.
- When I have the opportunity to determine on my own what I do in the Dutch class, I usual do a writing task.
- I write even if the teacher does not assign a writing task.
- I would like to spend more time on writing.
- Writing is a waste of time.
- I always look forward to writing lessons.
- I write because I have to at school.
- I like it to write down my thoughts.
- I would like to write more at school.

### Example: Questionnaire to Seventh' Graders Pleasure in Writing

- I love writing.
- Writing is my favorite school subject.
- When I write, I feel well.
- I hate writing.
- I write as soon as I get the chance.
- I make sure that I have to write as less as possible.
- I write more than my class mates.
- When I write, I prefer to do something different.
- Writing gives me pleasure.
- I just write, when I can get a good grade for it.
- Writing is boring.

- I like different kinds of writing.
- When I have the opportunity to determine on my own what I do in the Dutch class, I usual do a writing task.
- I write even if the teacher does not assign a writing task.
- I would like to spend more time on writing.
- Writing is a waste of time.
- I always look forward to writing lessons.
- ▶ I write because I have to at school.
- I like it to write down my thoughts.
- I would like to write more at school.

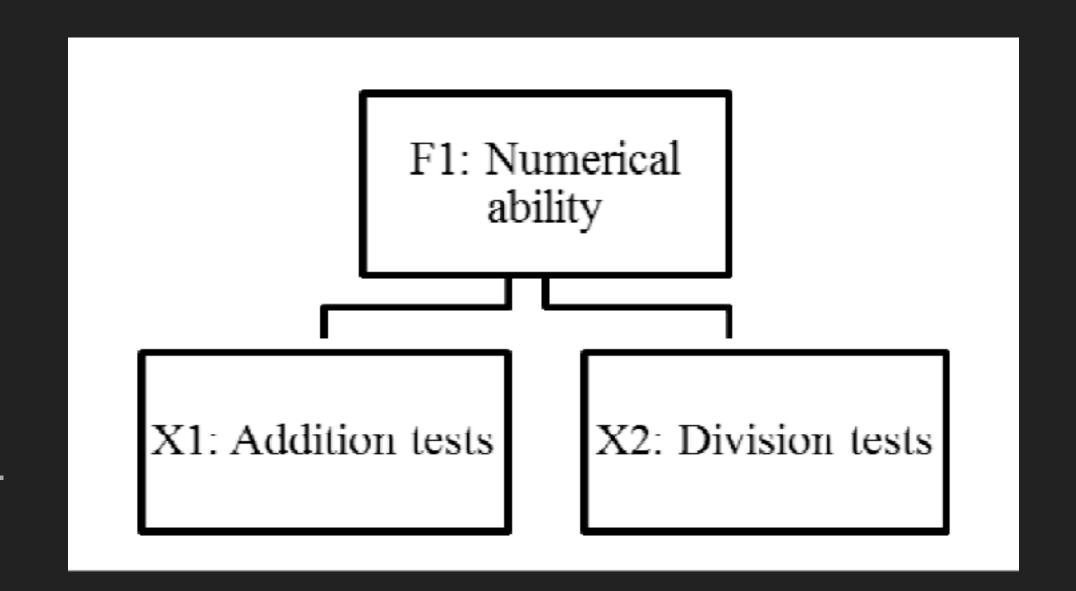
### Factor Analysis To Assess Construct Validity

### Factor analysis:

- A common data summarization technique.
- Used to regroup variables into a limited set of clusters based on shared variance.
- Helps to isolate constructs and concepts.

### In this context:

- Exploratory factor analysis detects the constructs i.e. factors that underlie a dataset based on the correlations between questionnaire items.
- The factors that explain the highest proportion of variance the questionnaire items share are expected to represent the underlying constructs.



### Factor Extraction

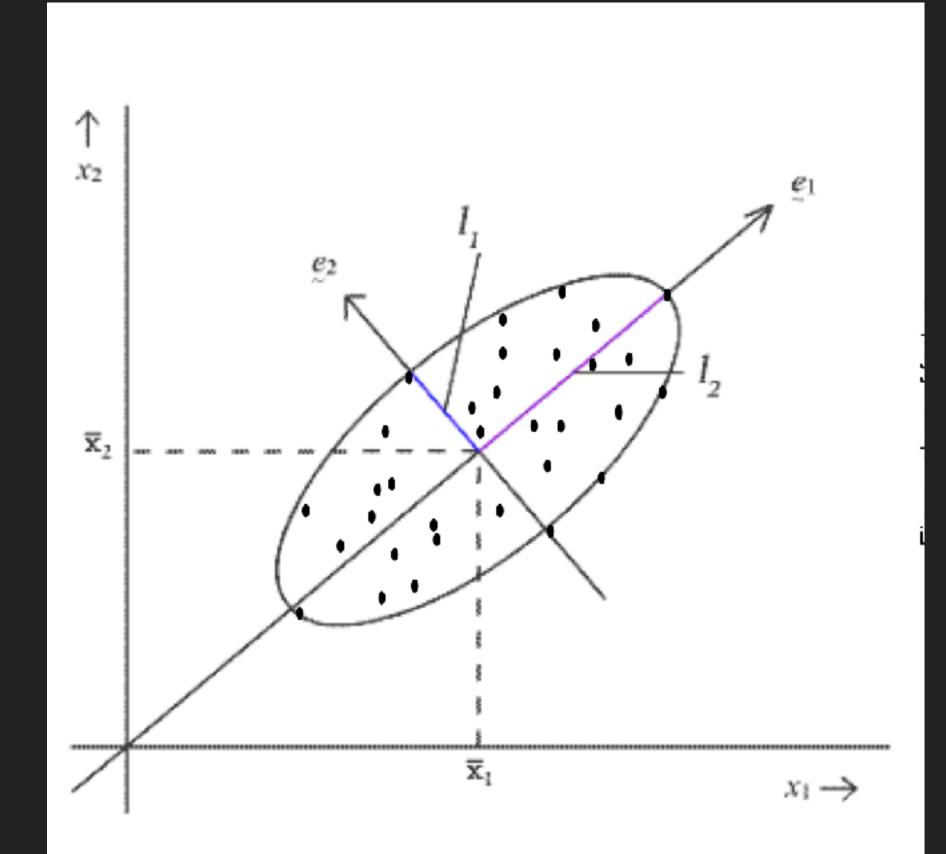


Figure 2. Scatterplot of two variables  $(x_1 \text{ and } x_2)$ . The lines  $e_1$  and  $e_2$  represent the eigenvectors of the correlation matrix of variables  $x_1$  and  $x_2$ . The eigenvalue of an eigenvector is the length of an eigenvector measured from one end of the oval to the other end.

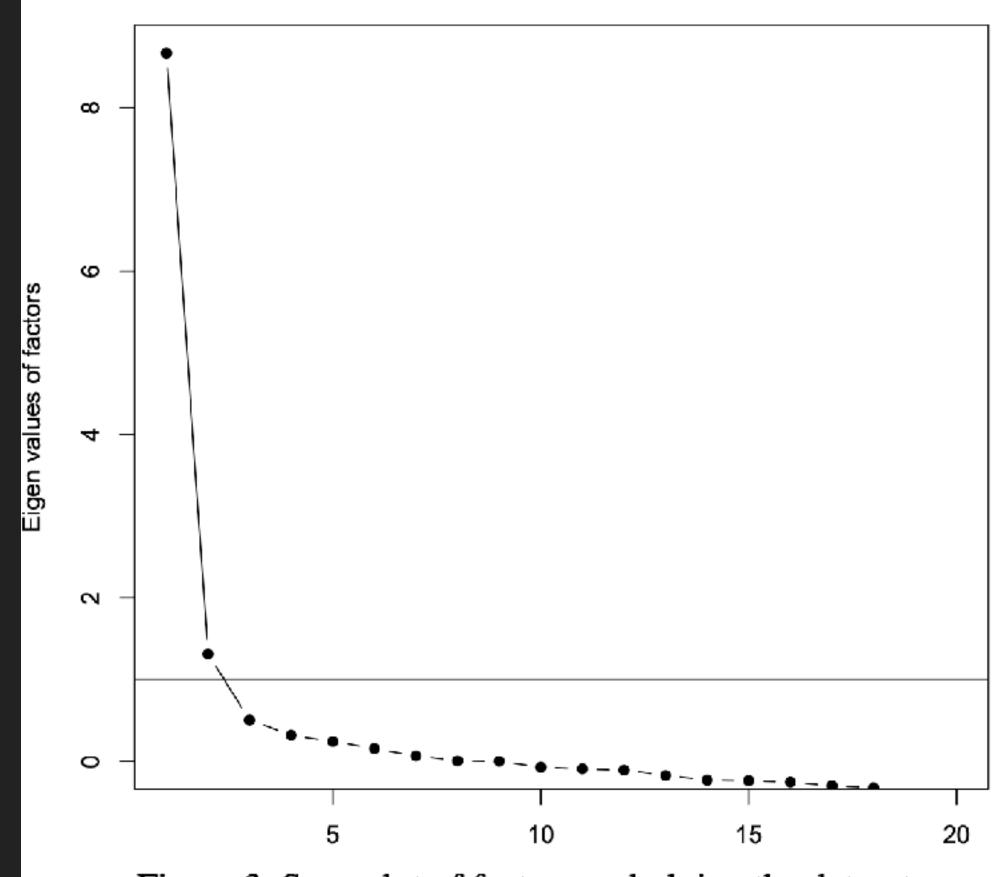


Figure 3. Screeplot of factors underlying the dataset. Every point represents one factor.

## The Positively Formulated Items in This Questionnaire Make Up the First Factor and the Negatively Formulated Items the Second Factor.

- p01. I love writing.
- p02. Writing is my favorite school subject.
- p03. When I write, I feel well.
- p04. I hate writing.
- > p05. I write as soon as I get the chance.
- p06. I make sure that I have to write as less as possible.
- p07. I write more than my class mates.
- p08. When I write, I prefer to do something different.
- > p09. Writing gives me pleasure.
- p10. I just write, when I can get a good grade for it.
- p11. Writing is boring.
- p12. I like different kinds of writing.
- p13. When I have the opportunity to determine on my own what I do in the Dutch class, I usual do a writing task.
- > p14. I write even if the teacher does not assign a writing task.
- p15. I would like to spend more time on writing.
- p16. Writing is a waste of time.
- p17. I always look forward to writing lessons.
- p18. I write because I have to at school.
- p19. I like it to write down my thoughts.
- p20. I would like to write more at school.

```
Call:
factanal(x = na.omit(passion), factors = 2, rotation = "oblimin")
Loadings:
     Factor1 Factor2
p01 0.547
                                              Factor1 Factor2
            -0.289
    0.747
p02
                             SS loadings
                                              6.141
                                                      3.534
    0.727
                                              0.307
p03
                             Proportion Var
                                                      0.177
                                              0.307
p04
             0.802
                             Cumulative Var
                                                      0.484
     0.625
p05
|p06
             0.641
p07 0.463
                             Factor Correlations:
p08 -0.133
             0.558
                                      Factor1 Factor2
     0.702
            -0.115
                                       1.000 -0.642
                             Factor1
p10
             0.597
                             Factor2 -0.642
p11
             0.680
     0.412
            -0.243
    0.719
                             Test of the hypothesis that 2 factors are sufficient.
             0.122
    0.739
                             The chi square statistic is 197.76 on 151 degrees of freedom.
p15
    0.758
                             The p-value is 0.00636
             0.771
p17
    0.917
p18
             0.739
p19 0.623
p20 0.771
```

## The Positively Formulated Items in This Questionnaire Make Up the First Factor and the Negatively Formulated Items the Second Factor.

- p01. I love writing.
- ▶ p02. Writing is my favorite school subject.
- ▶ p03. When I write, I feel well.
- p04. I hate writing.
- ▶ p05. I write as soon as I get the chance.
- > p06. I make sure that I have to write as less as possible.
- p07. I write more than my class mates.
- p08. When I write, I prefer to do something different.
- ▶ p09. Writing gives me pleasure.
- p10. I just write, when I can get a good grade for it.
- p11. Writing is boring.
- p12. I like different kinds of writing.
- ▶ p13. When I have the opportunity to determine on my own what I do in the Dutch class, I usual do a writing task.
- ▶ p14. I write even if the teacher does not assign a writing task.
- ▶ p15. I would like to spend more time on writing.
- p16. Writing is a waste of time.
- ▶ p17. I always look forward to writing lessons.
- p18. I write because I have to at school.
- ▶ p19. I like it to write down my thoughts.
- ▶ p20. I would like to write more at school.

```
Call:
factanal(x = na.omit(passion), factors = 2, rotation = "oblimin")
ll oadi nas :
     Factor1 Factor2
p01 0.547
                                               Factor1 Factor2
     0.747
p02
                              SS loadings
                                               6.141
                                                       3.534
     0.727
p03
                                               0.307
                              Proportion Var
                                                       0.177
p04
                                               0.307
             0.802
                              Cumulative Var
                                                       0.484
p05
     0.625
             0.641
     0.463
                              Factor Correlations:
p08 -0.133
             0.558
                                       Factor1 Factor2
     0.702
             -0.115
p10
             0.597
                                       -0.642
                              Factor2
p11
             0.680
     0.412
             -0.243
     0.719
             0.122
                              Test of the hypothesis that 2 factors are sufficient.
     0.739
                              The chi square statistic is 197.76 on 151 degrees of freedom.
p15
     0.758
                              The p-value is 0.00636
             0.771
     0.917
p17
p18
             0.739
     0.623
p20 0.771
```

Common pattern that reverse-phrased items load on a different factor (Schmitt & Stults, 1985)

### Cronbach's Alpha Analysis

```
Reliability analysis
Call: alpha(x = passion)
     alpha average_r mean
      0.93
              0.42
                      3.3
                            0.36
Reliability if an item is dropped:
                                                 Item statistics
          alpha
                                                               r.cor mean sd
                    average_r
          0.93
                    0.41
                                                     114 0.77 0.76 3.5 1.19
|p01
          0.93
                    0.42
p02
                                                      114 0.68 0.67 3.9 1.07
                    0.41
p03
          0.93
                                                      114 0.77 0.76 3.6 1.06
          0.93
                    0.42
p04-
                                                     114 0.70 0.69 3.3 1.29
                    0.43
          0.93
                                                     113 0.57 0.55 3.8 0.90
p05
                    0.42
          0.93
p06-
                                                     114 0.61 0.59 3.0 1.19
p07
          0.93
                    0.43
                                                      114 0.55 0.52 3.4 0.82
                                                 p08- 114 0.63 0.60 2.6 1.18
          0.93
                    0.42
p08-
                    0.41
          0.93
p09
                                                      114 0.78 0.77 3.6 0.98
p10-
          0.93
                    0.43
                                                 p10- 114 0.59 0.57 2.5 1.23
                    0.42
p11-
          0.93
                                                     114 0.60 0.58 3.1 1.26
p12
                    0.42
          0.93
                                                     114 0.62 0.59 3.0 1.00
p13
                    0.42
          0.93
                                                      114 0.60
                                                               0.57 3.4 1.08
p14
          0.93
                    0.42
                                                      114 0.71 0.70 3.7 1.01
p15
                                                      114 0.71 0.70 3.8 0.95
                    0.42
          0.93
p16-
          0.93
                    0.42
                                                     113 0.63
                                                               0.62 3.4 1.12
p17
                    0.41
          0.93
                                                      113 0.80 0.80 3.8 0.85
p18-
                                                 p18- 114 0.62 0.61 2.6 1.17
          0.93
                    0.42
p19
          0.93
                    0.42
                                                 p19 114 0.67 0.66 3.2 1.06
p20
          0.93
                    0.41
                                                     114 0.77 0.76 3.8 0.94
```

Table 4. Output of a Cronbach's alpha analysis in R

### Example: Multi-Item Scale for Brainstorming

- 1. The group aimed to generate as many ideas as possible.
- 2. All ideas were welcome, no matter how unconventional they were.
- 3. The group tried to combine similar ideas into one.
- 4. The group aimed to build on the ideas generated.
- 5. Ideas were generated first individually, then discussed as a group.
- 6. An organizer or group leader facilitated brainstorming for my session/group.
- 7. Group members criticized ideas proposed during the group/session. (R)

Responses on 5-point Likert scale, asked to what extent the statement reflected the way their group decided to work, from "not at all" to "completely."

Analyzed inter-item reliability (Cronbach's  $\alpha$ ), had to drop #7

From: Filippova, A., Trainer, E., & Herbsleb, J. D. (2017). From diversity by numbers to diversity as process: supporting inclusiveness in software development teams with brainstorming. Paper presented at the International Conference on Software Engineering, Buenos Aires, Argentina.

### Survey Examples

Ecosystem survey

http://cmu.ca1.qualtrics.com/jfe/form/SV\_d4M66VwPlZYd5kh

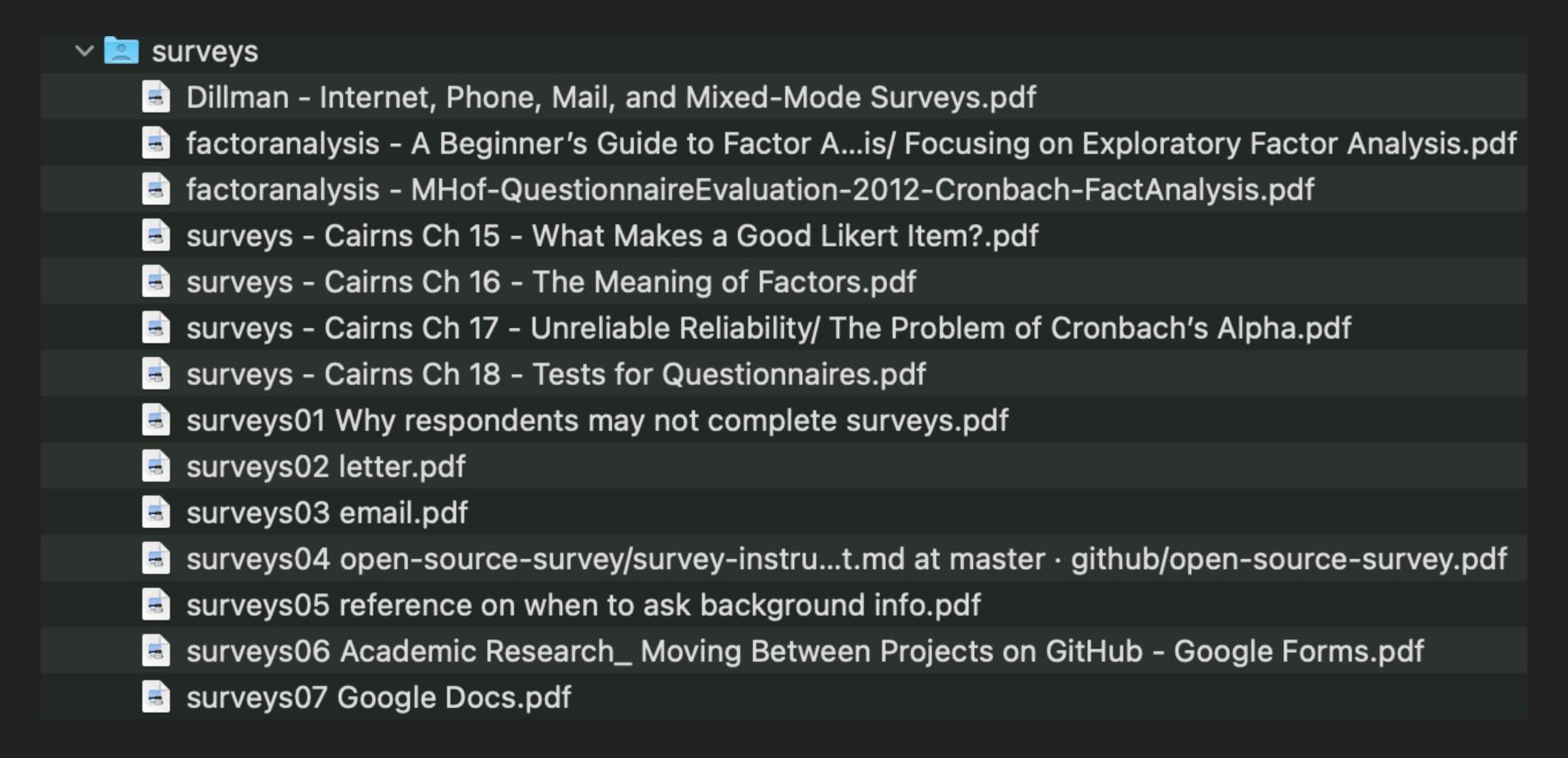
- Results: <a href="http://breakingapis.org/survey/">http://breakingapis.org/survey/</a>
- GitHub open source survey

https://github.com/github/open-source-survey

Results: <a href="http://opensourcesurvey.org/2017/">http://opensourcesurvey.org/2017/</a>

### Readings

https://drive.google.com/drive/folders/1ISOQIbw-cRmT47\_itpIJkTZua\_IASNgB?usp=sharing



### Credits

#### Graphics:

Dave DiCello photography (cover)

#### Content:

- Chapters from Dillman, D., Smyth, J. D., & Christian, L. M. (2014). Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method (4th ed.). Hoboken, NJ: Wiley.
  - ▶ Ch1: Sample Surveys in our Electronic World
  - ▶ Ch2: Reducing People's Reluctance to Respond to Surveys
  - Ch4: The Fundamentals of Writing Questions
  - ▶ Ch5: How to Write Open and Closed Ended Questions
- Hof, M. (2012). Questionnaire Evaluation with Factor Analysis and Cronbach's Alpha. Student project.
   Seminar in Methodology and Statistics. University of Groningen
- Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis.
   Tutorials in quantitative methods for psychology, 9(2), 79-94.
- Cairns, P. (2019). Doing better statistics in human-computer interaction. Cambridge University Press.
  - Ch15: What Makes a Good Likert Item?
  - Ch16: The Meaning of Factors
  - ▶ Ch17: Unreliable Reliability: The Problem of Cronbach's Alpha
  - Ch18: Tests for Questionnaires