Data Quality: Does Time of Semester Matter?

Linden Hillhouse and Ginette Blackhart

East Tennessee State University

DEPARTMENT of PSYCHOLOGY
College of Arts & Sciences
EAST TENNESSEE STATE UNIVERSITY

Introduction

- Previous research found that underpowered samples can increase Type II error rates (Vadillo et al., 2016).
- One possible source of underpowered samples is poor data quality.
- When many participants respond carelessly, data quality can be negatively affected (Meade & Craig, 2012).
- Meade and Craig (2012) suggest that online, non-laboratory studies run a greater risk of careless responding as the setting cannot be controlled and distractions may be prevalent.

Method

- Participants recruited through Sona during the first 2 weeks of Spring 2019, Fall 2019, & Spring 2020 semesters (Group 1) and the last 2 weeks of Fall 2018, Spring 2019, & Fall 2019 semesters (Group 2).
  - Group 1: $n = 138$, $M_{age} = 19.64$, $SD = 3.37$, 78% female
  - Group 2: $n = 779$, $M_{age} = 20.13$, $SD = 3.55$, 62% female
- Participants asked to recall and write about an embarrassing situation and then to complete several questionnaires assessing thoughts and feelings about the event, personality traits, and participant engagement in an online survey.
- 12 attention check items were embedded throughout the survey (Meade & Craig, 2012).
- Data quality was assessed using previously validated methods, including:
  - Time spent on survey;
  - Number of missed items;
  - Number of incorrect attention-check items;
  - Length of responses on two open-ended questions;
  - Self-reported diligence, interest, effort, attention;
  - Response bias
  - Internal consistency (Cronbach's α)

Results

- Group 1 had fewer missed items; longer responses to open-ended questions; reported higher levels of diligence, interest, effort, attention, and conscientiousness; and showed lower levels of response bias than Group 2.
  - When conscientiousness was included as a covariate in analyses, differences in number of missed items became non-significant; all other differences remained significant
- Only 46.13% of participants did not miss any attention check items.

Data Quality Indicator	Group	*Mean*	*SD*	*t*	*p*	*d*
Duration (min.)	1	28.90	13.37	.258	.80	.024
	2	28.56	14.40			
Missed Items	1	.38	1.40	-3.64	.24	-.19
	2	1.98	11.73			
Incorrect Attention Check Items	1	1.14	1.51	-1.84	.067	-.15
	2	1.41	2.04			
Length of Open-Ended Responses (# of words)	**1**	**46.66**	**16.66**	**6.22**	**<.001**	**.55**
	2	**36.93**	**18.47**			
Self-Reported Diligence	**1**	**57.10**	**5.49**	**7.72**	**<.001**	**.59**
	2	**52.73**	**8.97**			
Self-Reported Interest	**1**	**30.90**	**6.64**	**4.36**	**<.001**	**.40**
	2	**27.97**	**7.36**			
Self-Reported Effort	**1**	**4.44**	**.63**	**4.02**	**<.001**	**.34**
	2	**4.20**	**.83**			
Self-Reported Attention	**1**	**4.58**	**.55**	**5.12**	**<.001**	**.41**
	2	**4.30**	**.78**			
Response Bias Average	**1**	**4.18**	**.92**	**-5.02**	**<.001**	**-.36**
	2	**4.69**	**1.81**			
Response Bias Maximum	**1**	**11.95**	**3.50**	**-4.66**	**<.001**	**-.36**
	2	**13.64**	**5.76**			
Cronbach's Alphas	1	.83	.08	.17	.87	.06
	2	.824	.09			
Personality Factor	**Group**	***Mean***	***SD***	***t***	***p***	***d***
Conscientiousness	**1**	**33.86**	**5.06**	**-5.04**	**<.001**	**.47**
	2	**31.29**	**5.61**			

Conclusion

- Our results suggest that data collected at the end of the semester may be of lower quality than data collected at the beginning of the semester.
- This is perhaps because students may experience more time pressures and constraints at the end of the semester, perhaps encouraging students to multi-task while completing online surveys or to rush through the survey in order to receive research credits as quickly as possible.
- These results have implications for researchers in choosing when to collect data for online studies.
  - Knowing that data may vary in quality based on the time of semester collected, researchers may want to statistically control for this variation or collect data only during certain times of the semester.

Limitations / Future Research

- Given the high rate of missed attention check items, future research should determine ways to improve participant attention and engagement.
  - Variations in survey length and format may affect participant attention (e.g., Galesic & Bosnjak, 2009).
- Future research should examine time of semester effects across the entire semester, not just at the beginning and end of the semester.

Data quality may be worse at the end of the semester.

Acknowledgements

- Thank you to Dr. Blackhart for her guidance and support in the development of this research.
- Thank you to ETSU Honors College Office of Undergraduate Research & Creative Activities and to the ETSU Department of Psychology for funding my travel to SPSP.

Presented at the Society for Personality & Social Psychology, New Orleans, LA, Feb. 27-29, 2020. Questions about this research or interested in collaborating? Contact Dr. Ginni Blackhart at blackhar@etsu.edu.