



Nonresponse and Measurement Bias in Web surveys

Presented at ESRA 2015
Reykjavik, Island
July 14th, 2015

Background

- Web survey response rates are generally lower than response rates of self-administered surveys such as mail surveys (Couper, 2001).
- Substantial nonresponse has the potential to induce nonresponse bias.
- Web survey researchers have established various strategies to reduce nonresponse rates: appealing invitation e-mails, e-mail reminders, incentives etc. (Fan & Yan, 2010).
- However, incorporating nonrespondents into Web surveys does not necessarily reduce nonresponse bias and might cause additional measurement error.
- Nonresponse is caused by initial nonrespondents and survey breakoffs/ drop-outs (Bosnjak, 2001).

Research Questions

- Should survey break-offs be incorporated into Web surveys?
 - Do survey break-offs induce nonresponse bias?
 - Does convincing survey breakoffs to participate in Web surveys cause additional measurement bias?

- Do initial nonrespondents induce nonresponse bias?

- Do initial nonrespondents or survey break-offs cause the nonresponse bias to a greater extent?

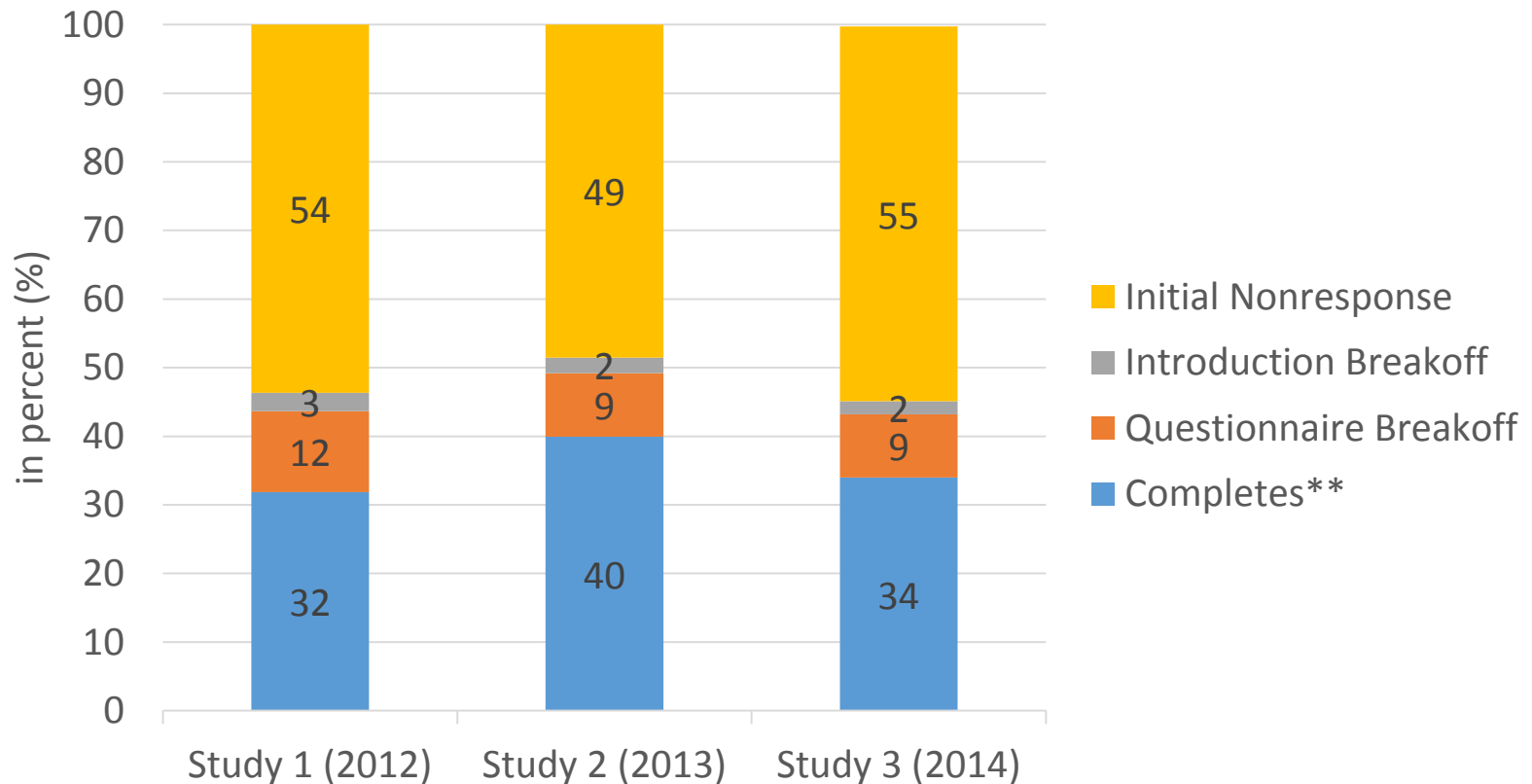
Data Basis

Studies of University Applicants

	Study 1	Study 2	Study 3
Sample:	university applicants (n=5,977)	university applicants (n=7,395)	university applicants (n=5,996)
Field phase:	July/August 2012	July/August 2013	July/August 2014
Study Design:	Between subjects design with random assignment		
AAPOR RR6:	32%	40%	35%
No. of Questions:	42	37	42
No. of Pages:	38	35	39

Response & Nonresponse Rates

Study 1-3*



* Respondents were excluded who completed the questionnaire but answered less than 50 percent of all applicable survey questions.

** Respondents who answered 50 percent or more of all applicable survey questions but did not complete the questionnaire were counted as survey break-offs.

Nonresponse Bias

$$\text{RNB} = \frac{N (\text{nonrespondents})}{N (\text{total number})} \times (p (\text{respondents}) - p (\text{nonrespondents}))$$

N (nonrespondents) = the number of nonrespondents in the target population

N (total sample) = the total number in the target population

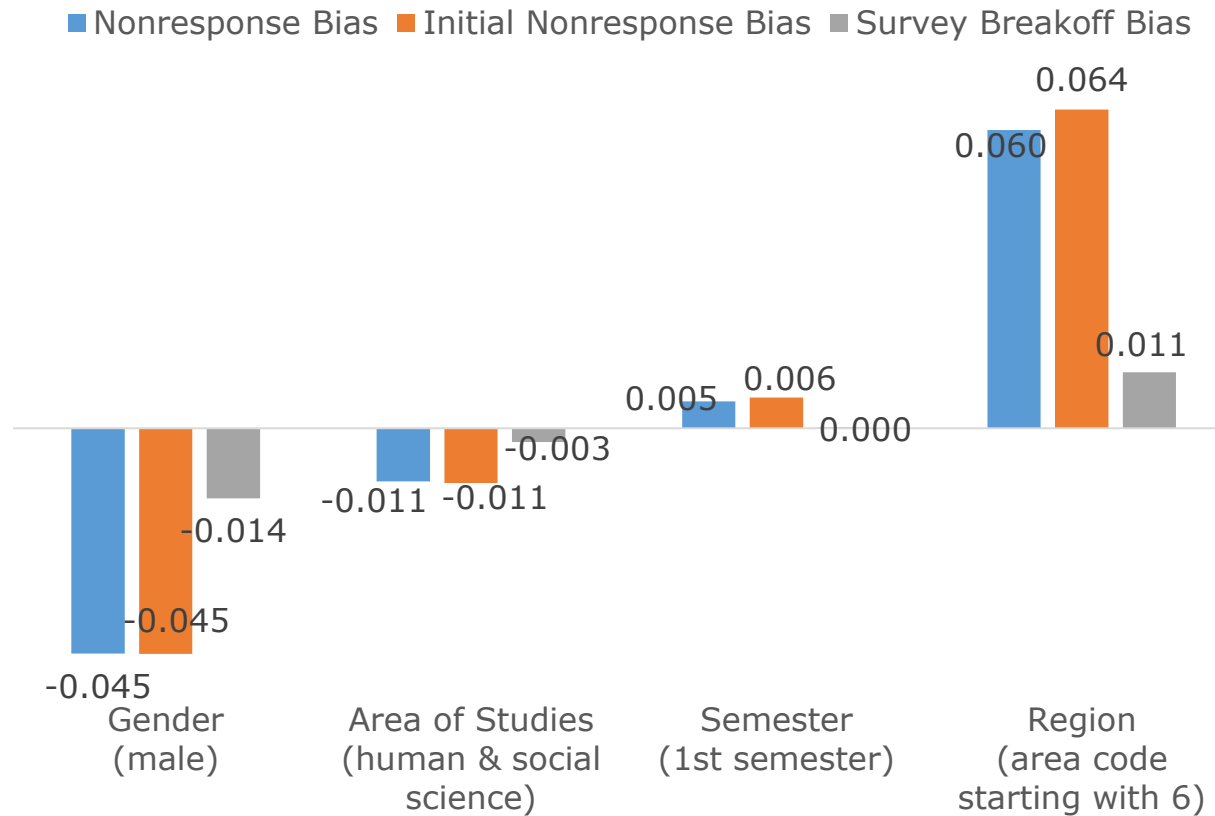
p (respondents) = the share of respondents in the target population

p (nonrespondents) = the share of nonrespondents in the target population

(Groves, 2006)

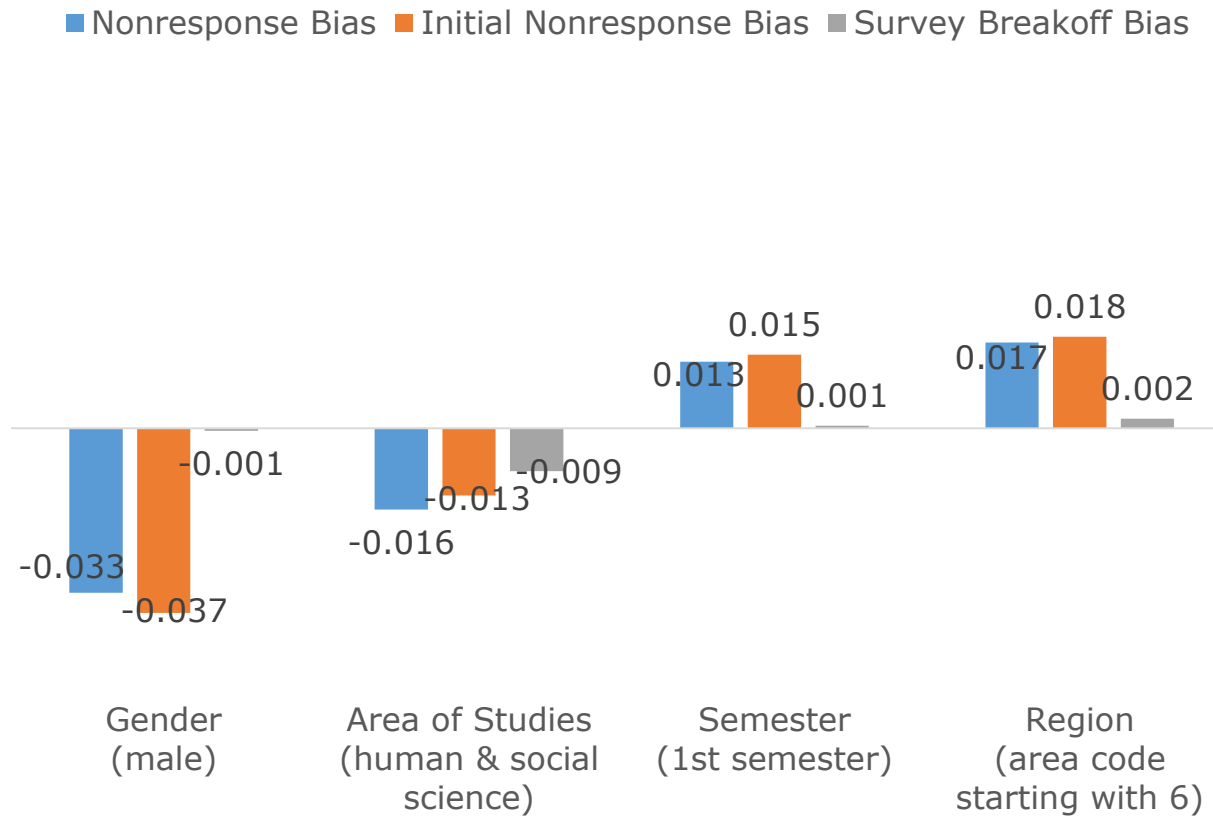
Nonresponse Bias

Study 1 (2012)



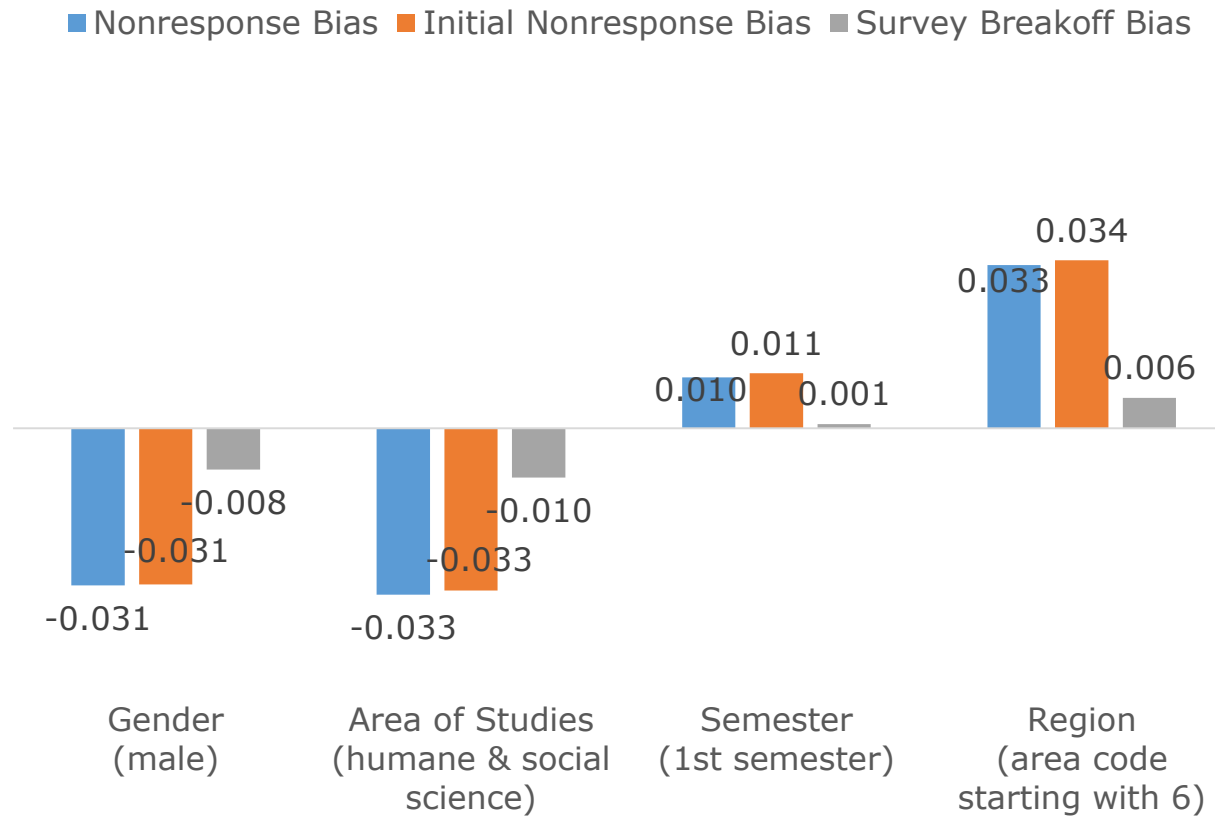
Nonresponse Bias

Study 2 (2013)



Nonresponse Bias

Study 3 (2014)



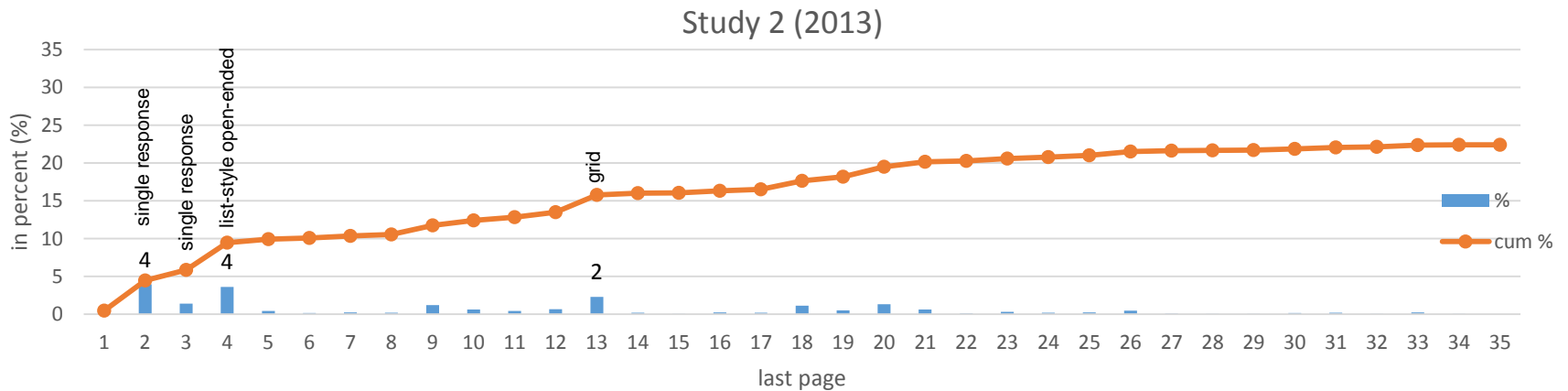
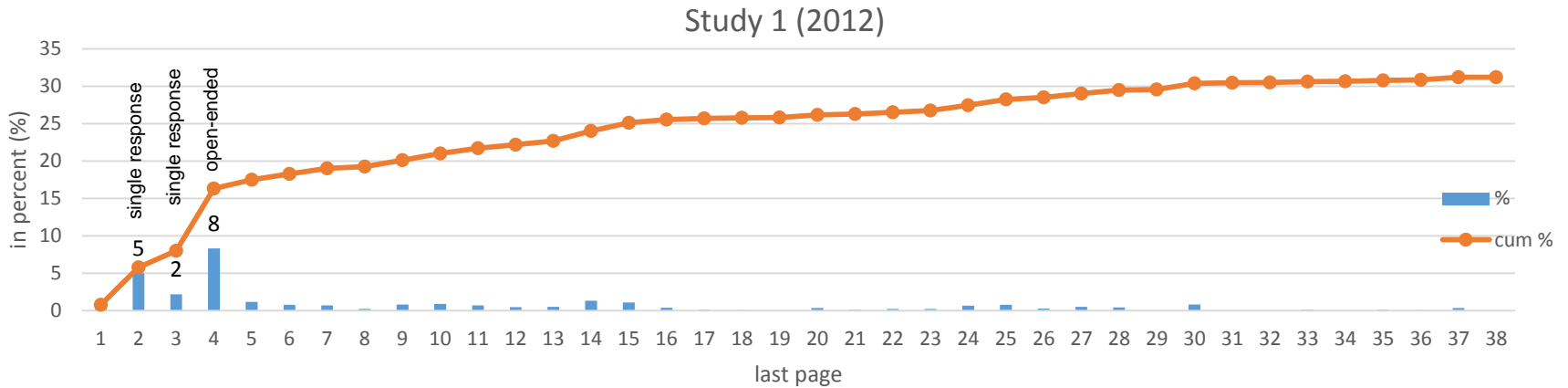
Summary

Results: Nonresponse Bias

- Bias estimates for all four variables were at a very low level.
- The initial nonresponse bias and the survey breakoff bias add to each other.
- In all studies the nonresponse bias for all variables is caused by the initial nonresponse bias to a greater extent.
- Therefore, it seems to be more important to reduce initial nonresponse.
- However, it is easier and more cost-effective to convince survey breakoffs to participate in Web surveys by optimizing the questionnaire design.

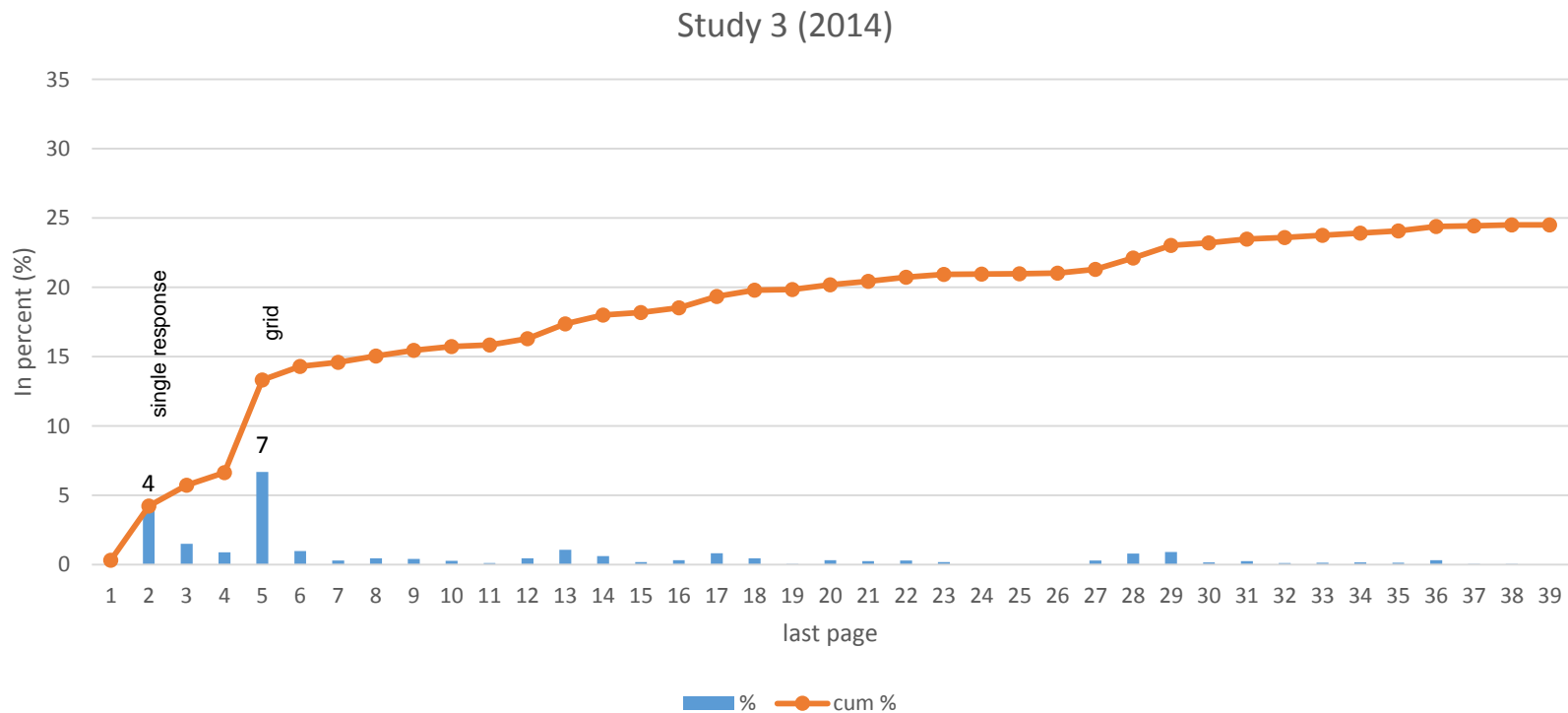
Breakoff by Page Number

Study 1 and 2



Breakoff by Page Number

Study 3



Measurement Bias (Response Time)

	Study 1		Study 2		Study 3	
	Beta	Sig	Beta	Sig.	Beta	Sig.
Semester (semester 2+)	19.9	.288	24.2	.000	7.2	.144
Region (area code ≠ 6)	2.0	.613	-14.1	.000	4.8	.000
Area of studies (engineering/science)	1.4	.719	6.1	.000	-4.8	.000
Gender (female)	2.7	.500	-13.5	.000	-1.0	.296
Breakoff (complete)	-77.1	.000	-13.7	.000	-9.3	.000
Constant	125.8	.000	44.3	.000	61.6	.000
R ²	.033		.049		.016	
Question type	narrative open-ended		list-style open-ended		grid	
Question position	4/38		4/35		5/39	

Note. Reference categories in brackets. Linear regression model.

Measurement Bias (Nondifferentiation)

	Study 1		Study 2		Study 3	
	Beta	Sig	Beta	Sig.	Beta	Sig.
Semester (semester 2+)	--	--	--	--	--	--
Region (area code ≠ 6)	--	--	--	--	--	--
Area of studies (engineering/science)	-.055	.000	-.001	.698	-.015	.581
Gender (female)	.053	.000	.003	.267	-.011	.692
Breakoff (complete)	.004	.636	-.015	.002	-.123	.000
Constant	.570	.000	.601	.000	.473	.000
R ²	.067		.002		.039	
Question type	grid		grid		grid	
Question position	15/38		13/35		5/39	

Note. Reference categories in brackets. Linear regression model.

Measurement Bias

(Number of characters of responses to open-ended questions)

	Study 1		Study 2		Study 3	
	Beta	Sig	Beta	Sig.	Beta	Sig.
Semester (semester 2+)	--	--	--	--	--	--
Region (area code ≠ 6)	2.5	.316	-9.2	.000	.6	.875
Area of studies (engineering/science)	4.4	.092	5.5	.023	-4.9	.164
Gender (female)	-10.4	.000	-15.3	.000	11.3	.002
Breakoff (complete)	-54.6	.000	-43.3	.000	-14.8	.028
Constant	146.5	.000	92.9	.000	76.2	.000
R ²	.044		.050		.013	
Question type	narrative open-ended		narrative open-ended		narrative open-ended	
Question position	4/38		9/35		7/39	

Note. Reference categories in brackets. Linear regression model.

Measurement Bias (Item Nonresponse)

	Study 1		Study 2		Study 3	
	Odds	Sig	Odds	Sig.	Beta	Sig.
Semester (semester 2+)	--	--	--	--	--	--
Region (area code ≠ 6)	1.2	.016	1.2	.000	.00	.913
Area of studies (engineering/science)	.8	.077	.9	.101	.05	.424
Gender (female)	1.0	.845	1.0	.556	.10	.156
Breakoff (complete)	5.5	.000	2.8	.000	6.53	.000
Constant	.2	.000	.5	.000	.04	.634
R ²	.138		.041		.525	
Question type	narrative open-ended		narrative open-ended		grid	
Question position	4/38		9/35		5/39	

Note. Reference categories in brackets. Linear regression model.

Summary

Results: Measurement Bias

- Breakoffs take significantly less time to answer questions.
- Breakoffs differentiate significantly less than completes.
- Breakoffs report shorter answers with a significantly lower number of characters than completes.
- Breakoffs are prone to leave items unanswered. They show significant higher item nonresponse rates than completes.



Incorporating breakoffs will increase the measurement error.

Conclusion

- Convincing survey break-offs to participate in Web surveys would increase response rates but reduce nonresponse bias only to a small extent.
- Furthermore, incorporating break-offs causes additional measurement error.
- Thus, survey break-offs should not be incorporated into Web surveys.
- Further analyses also need to examine if the quality of estimates becomes better with or without including break-offs. → Mean Squared Error
- Convincing initial nonrespondents to participate in Web surveys would reduce nonresponse bias to a greater extent but further research is also necessary to analyze the data quality of initial nonrespondents.



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Thank you.

**Darmstadt University of Technology
Department 02
Institute of Sociology
Research Methods
Dipl.-Soz. Anke Metzler**

Erst-Ludwig-Str. 3/ Im Carree
64283 Darmstadt
Germany

Phone: +49 6151/16-70972
Fax: +49 6151/16-72070

metzler@ifs.tu-darmstadt.de
