UNIWERSYTET JAGIELLOŃSKI
W KRAKOWIE

UNIWERSYTET JAGIELLOŃSKI
W KRAKOWIE

Controling and validating methods to control for response biases in self-report educational data.

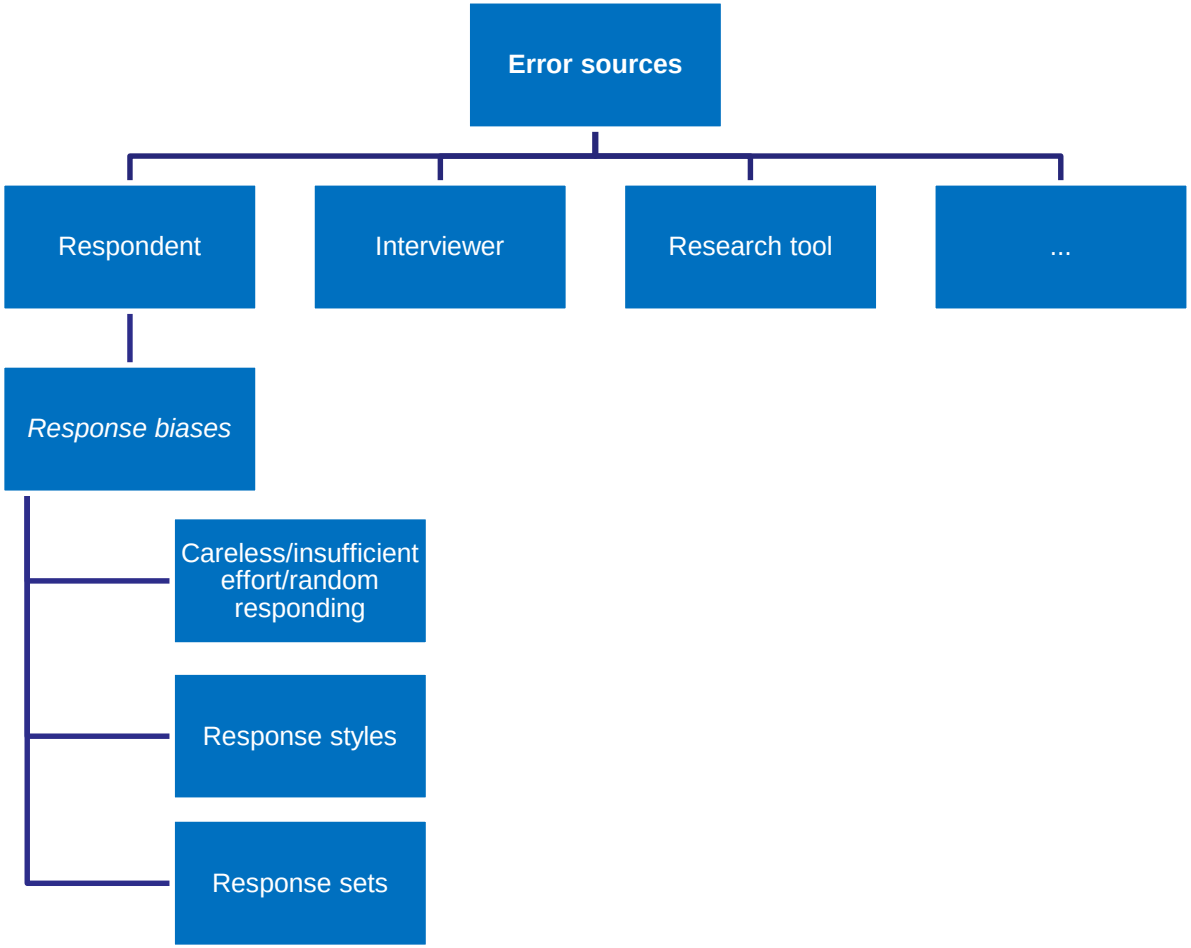
(work in progress)

Marek Muszyński
Institute of Sociology, Jagiellonian University

1. Presentation aim

- Present methods to control for response biases in self-descriptory data in educational studies.
- Comment on response biases risk in large-scale assessment studies (ILSAs) e.g. PIAAC, PISA, etc.

2. Potential measurement error sources.



2.1 What are response biases?

- systematical tendency to answer questions on other basis than they content (Paulhus, 1991)
- the research dates back at least to 1930s (e.g. Seashore, 1939) where „work methods" were first described and studied
- „"control" limited to ordinary instructions and demonstrations is incomplete, and that other unnoticed factors operate to modify the work method actually adopted." (Seashore, 1939)
- „Work methods may be temporary sets or may be habitual techniques of performance." (Cronbach, 1946)
- related or unrelated to item content (Nichols, Greene & Schmolck, 1989)
- most often: **interaction of respondent's traits and research context** (Baumgartner & Steenkamp, 2001; Mors, Kieruj & Vermunt, 2014; Schwarz, 1999; Weijters, 2006; Weijters et al., 2010; Ziegler & Buehner, 2009; Ziegler, 2015)

2.2 Response biases mechanisms- theoretical framework

- cognitive model of survey response (Tourangeau, Rips and Rasinski, 2000)
  - Comprehending
  - Retrieving
  - Integrating
  - Selecting/reporting
- interpretive visual heuristics (Tourangeau et al., 2004)
- response strategies (e.g. Silber et al., 2018)
- satisficing (Krosnick & Alwin, 1988; Krosnick, 1991, 1999)
  - hard
  - weak
- both frameworks are underinvestigated, but have gained some empirical confirmation (Gummer et al., 2018; Rossmann et al., 2017)

2.3 What response biases do we „know”?

- *careless responding, inattentive responding, random responding* (e.g. Meade i Craig, 2012)

- *response styles* (e.g. Weijters, 2006)

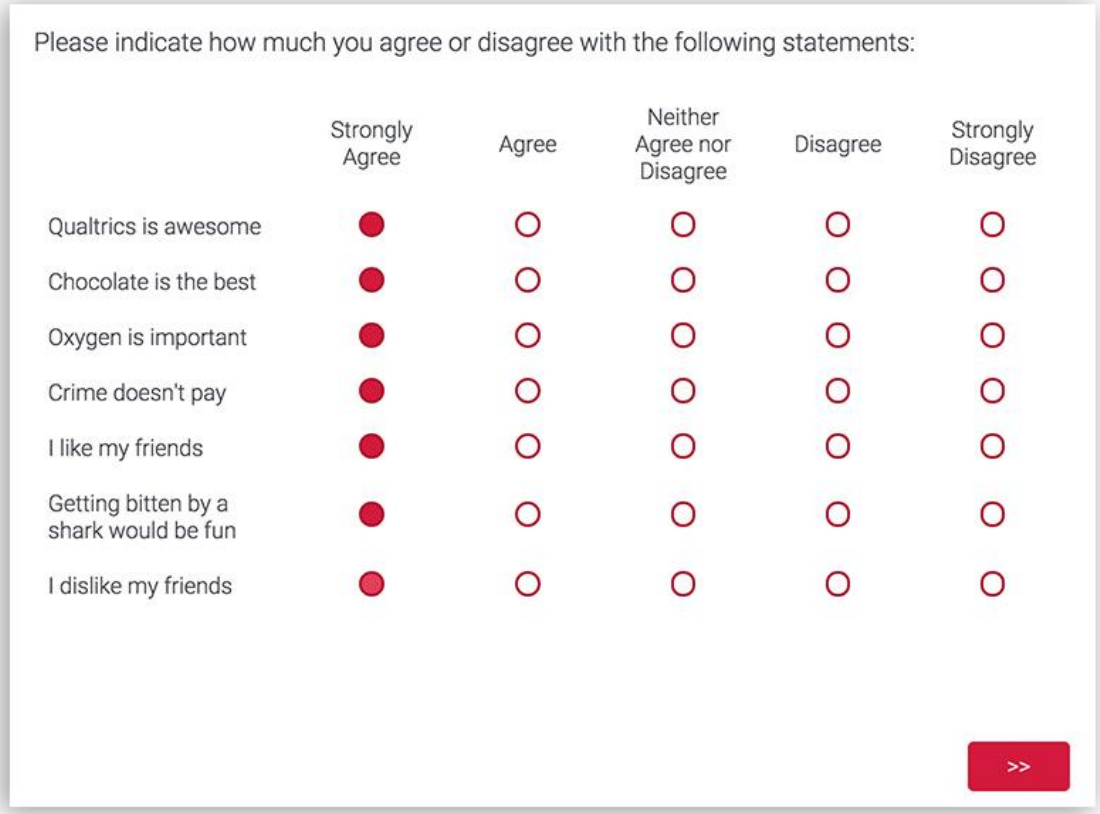
- *response sets* (e.g. Paulhus, 2002)

2.4 Response biases review

2.5 Careless responding

- *careless responding*, *inattentive responding, insufficient effort responding, random responding* (Curran, 2016; Fronczyk, 2014; Goldsmith i Clark, 2005; Meade i Craig, 2012; Meyvis, Oppenheimer i Davidenko, 2009; Osborne i Blanchard, 2010)
- only few % of respondents? (Johnson, 2005) or up to 50%? (Meyvis et al., 2009)
- limited knowledge on mechanisms of this effect (Curran, 2016; Meade & Craig, 2012)
- special focus in web surveys (Johnson, 2005)
  - accounbtability
  - anonimity
  - bots (e.g. MTurk)

2.6 Response biases review-response styles



2.7 Response styles

- overusing of a given response category (Khorramdel & van Davier, 2014; Mors, Kieruj & Vermunt, 2014)
- stable in time? (Jackson & Messick, 1958; Weijters, 2006)
- many response styles known:
- *extreme response style* (ERS)
- *middle response style* (MRS)
- *acquiescence response style* (ARS)
- *disacquiescence response style* (DARS)

(review: Van Vaerenbergh & Thomas, 2013)

2.8 Socially desirable responding

2.9 Socially desirable responding (SDR)-definitions

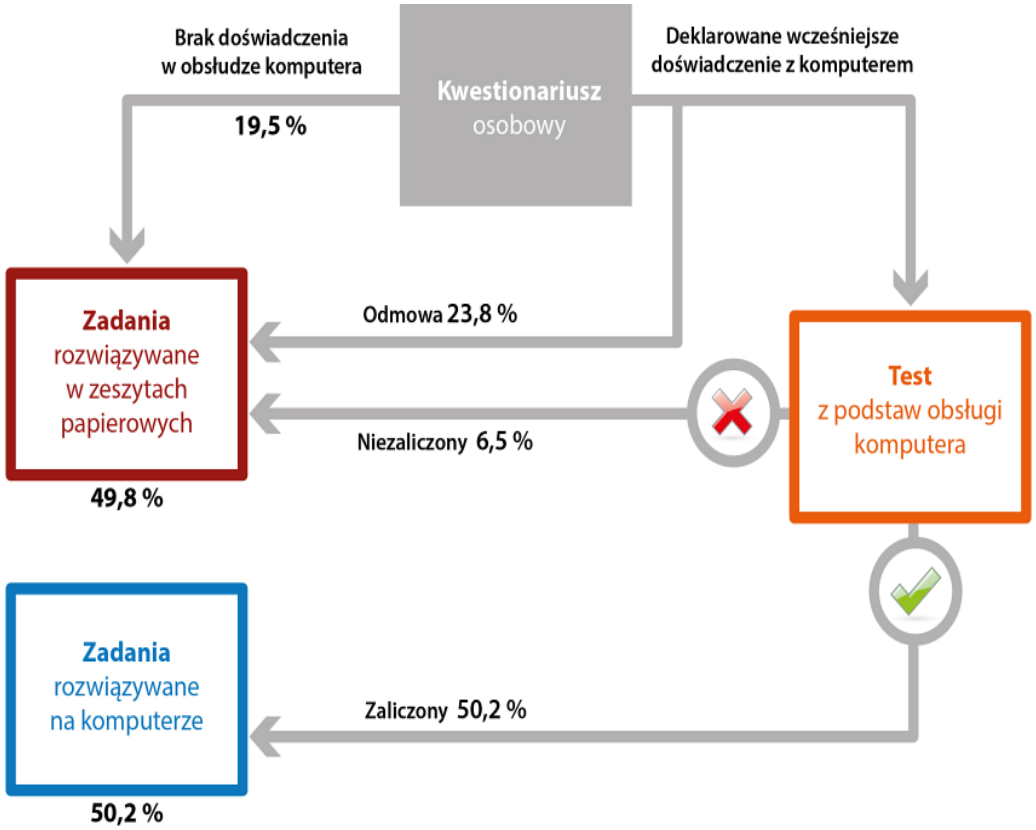
- „Social desirability is the tendency of some respondents to report an answer in a way they deem to be more socially acceptable than would be their "true" answer. ” (Callegaro, 2008)
- „Tendency to give (overly) positive self-descriptions” (Paulhus, 2002)
- „Participants’ tendency to describe themselves in favorable terms by adhering to socioculturally sanctioned norms” (De Jong et al., 2010)
- „Obtain approval by responding in a culturally appropriate and acceptable manner” (Crowne & Marlowe, 1960)
- **„general positive bias of self-perception”** (Wojciszke, 2010)
- „evolutionary mechanism” (Hogan, 2005)
- faking (Ziegler, McCann & Roberts, 2011)
- forensic psychology, psychopathology (Hildebrand et al., 2018)
- sensitive, intrusive, loaded questions (Krumpal, 2013; Tourangeau & Yan, 2007)

3.1 Bias example 1

- PISA 2012 (OECD, 2014; Pokropek, 2014)
- „Are you familiar with this mathematical concept”?

Concept	Know it well, understand the concept
Subjunctive scaling*	12,8%
Proper number*	50,4%
Declarative fraction*	29,6%
Polygon	88,9%

3.2 Bias example 2



In the PIAAC study around 20% of the participants that declared computer-proficiency refused to perform a simple computer-proficiency test. Another 6,5% failed this very simple test. These results point to overclaiming in the self-report data (Burski et al., 2013, PIAAC study, Polish sample)

3.3. Bias example 3

- students were asked about their predictions for their performance on a forthcoming final exam
- no one expected to score below the average (Taylor & Brown, 1988)
- what percent of Americans think their driving skills are better than average?
- 88% (Svenson, 1981)
- what percent of university teachers think they possess better than average teaching skills?
- 94% (Cross, 1997)

3.4 Bias example 4

- *„In a survey conducted by experts from the Faculty of Physics of the University of Poznań about 80-90% of the respondents declared excellent or satisfying sight. Only 20% passed an objective measurement test.”*

(http://www.biznes.newseria.pl/news/dobra-kondycja-rynku,p757362557)

3.5 Bias-comment

- general positive bias of self-perception/socially desirable responding also present in:

- ❑ low-stakes research
- ❑ no incentives
- ❑ non-intrusive questions

(http://www.biznes.newseria.pl/news/dobra-kondycja-rynku,p757362557)

4.1 Careless responding- sources

- low motivation
- low ability (e.g. cognitive skills, literacy skills)
- research procedure (e.g. cognitive burden, difficulty, length)
- (social) context

(Meade & Craig, 2012)

- A critical problem in web-based research?

(Fleischer et al., 2015; Gummer et al., 2018; Johnson, 2005)

  - anonymity
  - accountability
  - ease of responding, lower seriousness
  - bots

4.2 Response styles- sources

- low motivation
- low ability (e.g. cognitive skills, literacy skills)
- research procedure (e.g. cognitive burden, difficulty, length)
- (social) context
- personality
- culture

(He & van de Vijver, 2013, 2015; Van Vaerenbergh i Thomas, 2013)

4.3 Biases of self-perception- mechanisms and sources

- **intentional**
- faking
- avoiding answer
- impression management
- (Krumpal, 2013; Paulhus, 2002; 2011; Ziegler, 2011)
- **unintentional**
- universal cognitive biases (better than average, unrealistic optimism, overconfidence effects; Burrus et al., 2011; Wojciszke, 2010)
- protection of self-esteem; biased self-perception; low self-knowledge; lack of insight (Alicke, 2011; Ellison et al., 2006; Snyder, 1974; Zaborowski, 1989; Wojciszke, 2010)
- psychpathology (Crowne & Marlowe, 1960)
- **interaction of person and situation (Ziegler, 2015)**

5.1 Response biases consequences

- validity threat (Holden, 2007)
- systematic measurement error (CIV; *spurious variance*; Schmidt, Le & Ilies, 2003; Ziegler et al., 2011)
- higher nonresponse, missing data
- distorted means and variances, „theta shift”
- skewed distributions, misrepresented factorial structure
- distorted internal consistency
- changed multivariate correlations
- obscured cultural differences

(Khorramdel & Van Davier, 2014; Johnson & Van de Vijver, 2002; Krosnick & Alwin, 1988; Maniaci & Rogge, 2014; Paulhus, 1991; Pokropek, 2014).

5.2 Response biases consequences

- protocol invalidity
- infringement of measurement assumption
- inferential errors
- uninterpretable data

(Johnson, 2005; Kurtz, 2001; Maniaci & Rogge, 2014)

6. Call for research-research rationale

Need of:

- a method:
  - cost-effective,
  - easy-to-use,
  - valid,
  - *a posteriori* (*post hoc*),
  - flexible: different scales, modes, contexts, populations
- more comparison, validation and metanalytic studies
- **problem of response biases in low-stakes research**
- **very few research out of the „trinity" context: faking, sensitive questions, forensics/diagnostics** (seldom researched topic in educational studies and self-assessment of skills)
- **methods used in the above fields most often not applicable in large-scale surveys** (Ziegler et al., 2011)
- **many response patterns seem aberrant- but are they really are?**

7.1 Overclaiming technique

- dated back at least to 1970s and linguistics- word recognition research (Zimmerman et al., 1977)
- in social sciences (Paulhus et al., 2003): based on rating familiarity with a list of objects, among which we intertwine **reals** and **foils**

Format of the Over-Claiming Questionnaire (OCQ)

Using the following scale as a guideline, write a number from 0 to 6 beside each item to indicate how familiar you are with it

Never heard of it — Very familiar

0 1 2 3 4 5 6

Physical Sciences

_____ Manhattan Project	_____ asteroid	_____ nuclear fusion
_____ cholarine	_____ atomic number	_____ hydroponics
_____ alloy	_____ plate tectonics	_____ photon
_____ ultra-lipid	_____ centripetal force	_____ plates of parallax
_____ nebula	_____ particle accelerator	_____ satellite

Note. Of the 15 items above, the following 3 are foils: cholarine, ultra-lipid, and plates of parallax. Other topic categories include literature, art, history, social science, language, contemporary culture, and consumer products.

7.2 Research

- PISA 2012 data (OECD, 2014)
- familiarity with mathematical concepts (13 items) + questionnaire
- overclaim measure (3 items-"foils") (Paulhus et al., 2003; Vonkova et al., 2018)
- 2881 respondents (missing data removed)

Item	Thinking about mathematical concepts: how familiar are you with the following terms?
ST62Q01	a) Exponential Function
ST62Q02	b) Divisor
ST62Q03	c) Quadratic Function
ST62Q06	e) Linear Equation
ST62Q07	f) Vectors
ST62Q08	g) Complex Number
ST62Q09	h) Rational Number
ST62Q10	i) Radicals
ST62Q12	k) Polygon
ST62Q15	m) Congruent Figure
ST62Q16	n) Cosine
ST62Q17	o) Arithmetic Mean
ST62Q19	p) Probability
	Foils used for signal detection adjustment
ST62Q04	d) <Proper Number>
ST62Q11	j) <Subjunctive Scaling>
ST62Q13	l) <Declarative Fraction>

7.3 Research

- Is overclaiming technique a valid measure of controlling for response biases in low-stakes data? (very mixed results in other fields: Paulhus, 2011)
- What other response biases are present in this data? (careless responding?; Kam et al., 2015)
- How to control them? (Meade & Craig, 2012)
- What contextual variables are related to response biases?

- ❑ Prediction: OCT will act as a classical suppressor for the math familiarity-mathi ability relation
- ❑ Prediction: OCT relations with math ability will be mediated by careless/inattentive responding
- ❑ Prediction: response biases presence in the self-report will correlate negatively with math ability
- ❑ Prediction: male will yield lower data quality; low escs will yield lower data quality

6.2 Research

	Model 1		Model 2		Model 3		Model 4		Model 5	
	B	p>z	B	p>z	B	p>z	B	p>z	B	p>z
math familiarity	38,8	***	-	-	42,41	***	40,67	***	36,29	***
overclaiming	-	-	-3,21	*	-13,80	***	-9,67	***	-6,82	***
within-person SD	-	-	-	-	-	-	7,68	***	13,12	***
Mahalanobis distance	-	-	-	-	-	-	-	-	-11,07	***
R2 Level1		0,26		0,01		0,30		0,30		0,30
R2 Level2		0,36		0,01		0,44		0,43		0,40

- Multilevel regression: plausible values from matb ability as dependent variable
- Math familiarity- pv from GRM model
- Overclaim- pv from GRM model (uirt, Kondratek, 2016)
- Within-person standard deviation (egen rowsd command)
- Mahalanobis distance (R „careless" package; Yentes & Wilhelm, 2018)
- School- and participant-level weights used
- Snijders&Bosker (pseudo)R2 used (mltrsq, Stata 14)

6.3 Research-path analysis

	Direct effects		Indirect effects		Total effects	
	B	p>z	B	p>z	B	p>z
math ability <-						
within_sd	15,05	***	-	-	15,05	***
Mahalanobis distance	-12,84	***	-	-	-12,84	***
math familiarity	41,17	***	**6,72**	***	47,89	***
overclaiming	-10,63	***	**-7,90**	***	-18,53	***
Mahalanobis distance <-						
math familiarity	-0,29	***	-	-	-0,29	***
overclaiming	0	ns	-	-	0	ns
within_sd <-						
math familiarity	0,20	***	-	-	0,20	***
overclaiming	-0,53	***	-	-	-0,53	***

6.4 Research-path analysis

	Gender	p>z	ESCS	p>z
math familiarity	-0,19	***	0,29	***
overclaiming	0,15	***	-0,04	*
within_sd	-0,27	***	0,10	***
Mahalanobis distance	0,21	***	-0,08	***

7. Brief sum up

- **overclaiming technique acts as classical suppression in self-report math familiarity and math ability relation**
- overclaiming relations with math ability is mediated through careless responding indices
- **(**a lot of overclaimers are just straightliners!)
- **male:** higher overclaiming, lower data quality
- socio-economic status: slightly higher data quality for higher status participants

8. Discussion

UNIWERSYTET JAGIELLOŃSKI
W KRAKOWIE

Thank you for your attention!
marek.a.muszynski@gmail.com

Marek Muszyński
Institute of Sociology, Jagiellonian University