

Focus on mobile surveys: do the number of scale points and scale order affect rating scale results?

March 16, 2017 (Berlin)
General Online Research 2017



Prof. Dr. Andreas Krämer

BiTS Business and Information
Technology School
Reiterweg 26b, 58636 Iserlohn
andreas.kraemer@bits-laureate.de

exeo Strategic Consulting AG
Wittelsbacherring 24, 53115 Bonn
andreas.kraemer@exeo-consulting.com
Mob: +49 (0) 178 256 2241

Agenda

- 1. Introduction: the rise of mobile**
 - 2. Research method: online testing**
 - 3. Results: three different studies**
 - 4. Discussion: risk when changing scales**
 - 5. Outlook: what are the learnings?**
-

Agenda

- 1. Introduction: the rise of mobile**
- 2. Research method: online testing**
- 3. Results: three different studies**
- 4. Discussion: risk when changing scales**
- 5. Outlook: what are the learnings?**

Why the usage of rating scales is relevant when designing mobile surveys ...

Relevance of mobile surveys and rating scales

Emerging methods in use

	Rank	Labels	In Use
Mainstream	1	Mobile Surveys	75%
	2	Online Communities	59%
Wide Adoption	3	Social Media Analytics	52%
	4	Text Analytics	46%
	5	Webcam-Based Interviews	43%
	6	Mobile Qualitative	42%
	7	Big Data Analytics	38%
	8	Micro-surveys	35%
	9	Eye Tracking	35%
	10	Mobile Ethnography	33%
	11	Behavioral Economics Models	29%
Niche	12	Research Gamification	25%
	13	Prediction Markets	24%
	14	Facial analysis	24%
	15	Crowdsourcing	16%
	16	Neuromarketing	16%
	17	Virtual Environments/Virtual Reality	14%

GRIT REPORT, QIII, 2016;

Typical rating scale mobile research

Mail 15:06 26%
x1.rogator.de

How satisfied are you with the following characteristics of the questionnaire:

	1 = very satisfied	2	3	4	5 = very dissatisfied
Topic of the interview	☐	☐	☐	☐	☐
Length of the interview	☐	☐	☐	☐	☒
Graphical elements (images)	☐	☐	☐	☐	☐
Clarity of the questions	☐	☐	☐	☐	☐
Interview overall	☐	☐	☐	☐	☐

Continue

- How sensitive are results when changing rating scales?
- How robust are test results (taking into account the relevance of mobile)?
- What are the recommendations?

Results of experiments to measure the influence of the scale format on the results are contradictory

Results on the impact of different rating scales on survey results

Number of scale points

No major impact of scale points

- **Leung (2011):**
„no major difference in internal structure in terms of means, standard deviations...”
- **Dawes (2012):**
“The study found that the 5- and 7-point scales produced the same mean score as each other, once they were rescaled.”

Major impact of scale points

- **Weijters, Cabooter & Schillewaert, (2010):**
“... We find evidence of strong effects of scale format on response distributions and misresponse to reversed items.”
- **Garland, (1991):**
„the presence or absence of a mid-point on an importance scale produces distortions in the results obtained...”

Order of rating scale

No major impact of scale order

- **Christian et al. (2009):**
“We find that consistently presenting the positive end of the scale first did not impact responses but increases response times.”
- **Maloshonok & Terentev (2016):**
„... resultes with ascending (from negative to positive) or descending (from positive to negative) order ...do not differ substantially.”

Major impact of scale order

- **Toepoel et al. (2009):**
„Our results thus provide empirical support ... for the theory of satisficing and primacy effects”.
- **Yan & Keusch (2015):**
„... mean ratings are shifted toward the starting point of the scale, consistent with the prediction of satisficing and the anchoring-and-adjustment heuristics”

Agenda

1. Introduction: the rise of mobile

2. Research method: online testing

3. Results: three different studies

4. Discussion: risk when changing scales

5. Outlook: what are the learnings?

In order to measure the effects of different scale designs on survey results, three different studies were used

Study profile (scale testing)

Scale test # 1

- Online survey
- Field: August 2016 (Online access panel)
- Topic: Satisfaction with the interview (details + overall)
- Test: 2*2 (scale points 5/6; pos.-neg./neg.-pos.)
- Sample: n=586 (28 % mobile)
- Time: 14 minutes

Scale test # 2

- Online survey
- Field: January 2017 (customer contact data)
- Topic: Satisfaction with the interview (details + overall)
- Test: 2*2 (scale points 5/7; pos.-neg./neg.-pos.)
- Sample: n=3.022 (17 % mobile)
- Time: 3,5 minutes

Scale test # 3

- Online survey
- Field: March 2017 (Online access panel)
- Topic: Satisfaction with mobility offers
- Test: 2*2 (scale points 5/10; pos.-neg./neg.-pos.)
- Sample: n=520 (19 % mobile)
- Time: 15,8 minutes

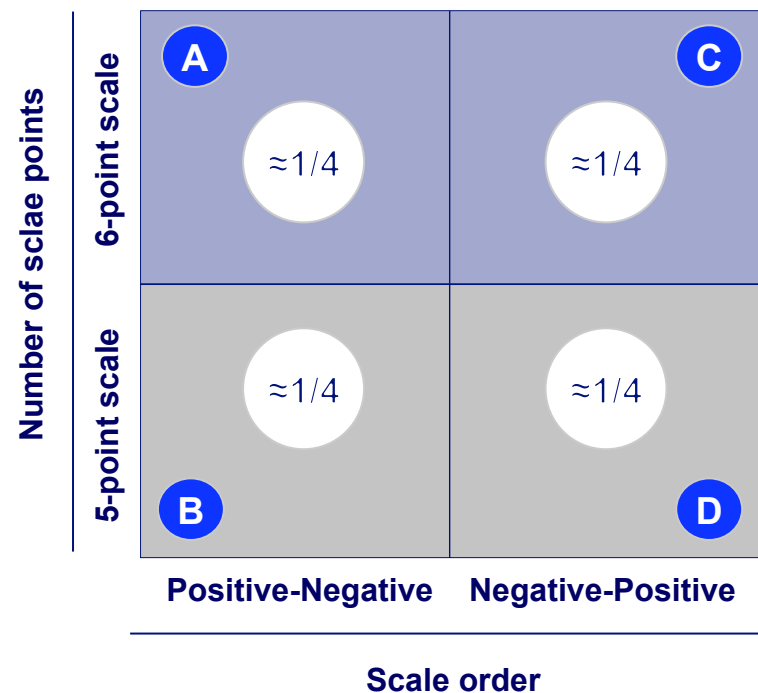
Field work and data management

- All three studies were conducted by Rogator AG, Nuremberg; studies #1 and #3 are part of the study “Pricing Lab” (co-operation of exeo and Rogator)

During the online interview a 2*2 factor design was used when rating the respondents' satisfaction regarding the interview

Study # 1

2 * 2 test design: number of scale points vs. scale order



Test environment

- 5 Statements at the end of the interview.
- Randomized test groups (no significant differences occurred in terms of age, income, device used for the interview and intensity of participation in online surveys; Chi-Square-test).
- Topic: Satisfaction with the interview (details + overall rating).
- Framing: During the interview a 6-point scale was used (pos.-neg.).
- The interview was rather long (14 min.).

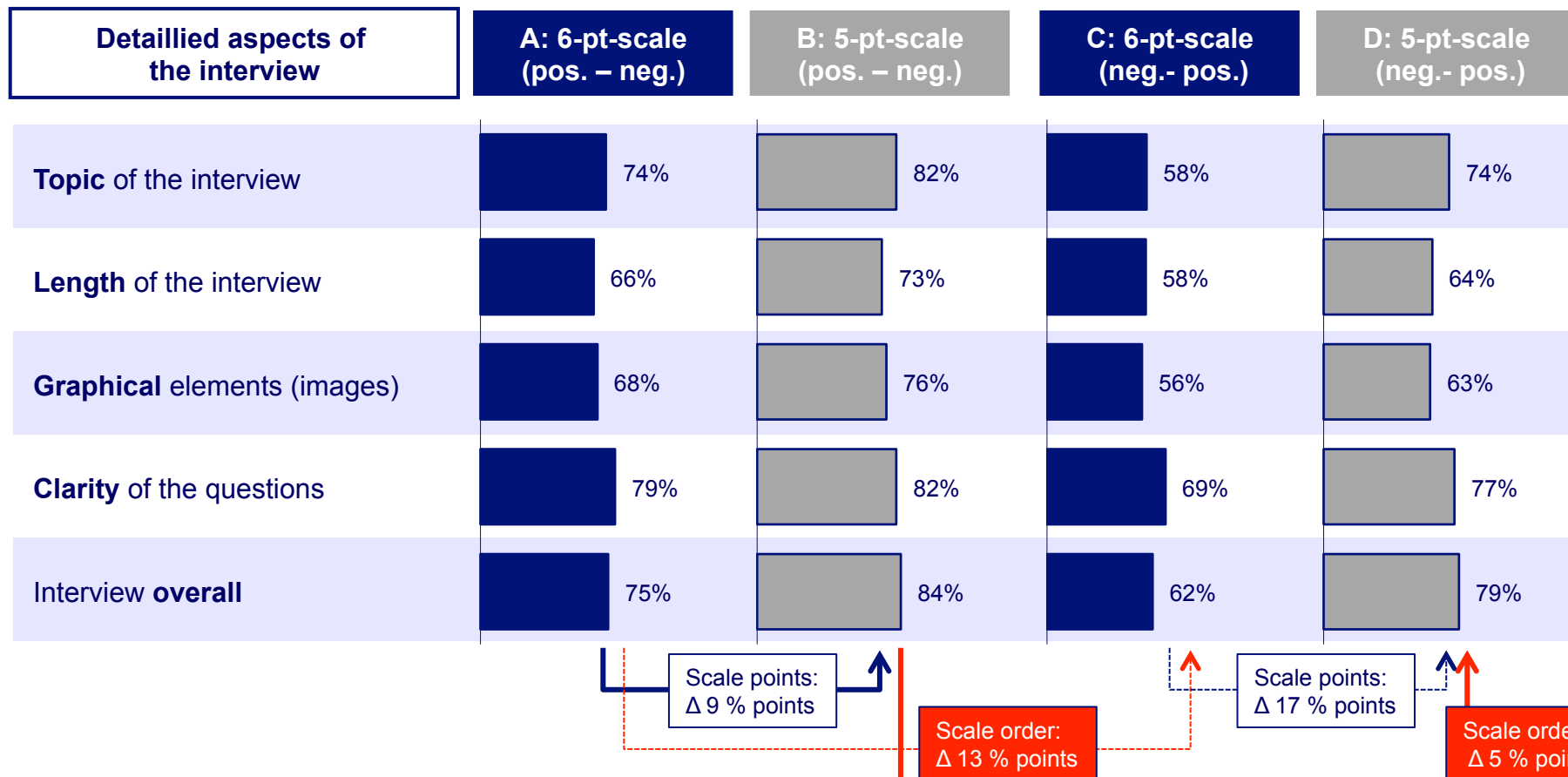
Agenda

1. Introduction: the rise of mobile
 2. Research method: online testing
 - 3. Results: three different studies**
 4. Discussion: risk when changing scales
 5. Outlook: what are the learnings?
-

Both, the number of scale points and scale order obviously influence the survey results (top-2 %)

Study # 1

Degree of satisfaction with the interview (% top-2 „satisfied / very satisfied“)¹⁾



1) How satisfied are you with the following characteristics of the questionnaire?

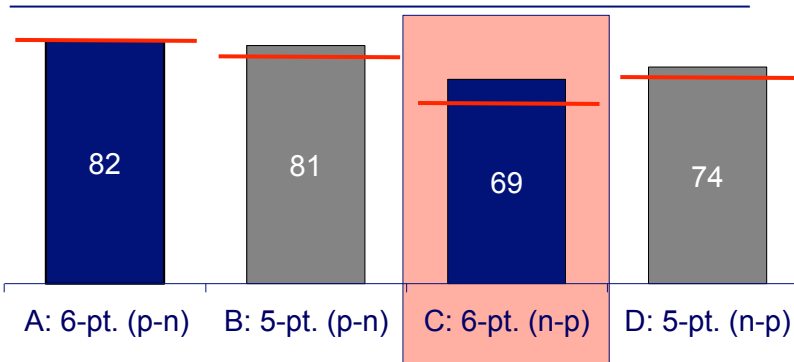
Source: **exeo** Strategic Consulting AG / Rogator AG

After standardization: Results still differ clearly across test groups – particularly low scores for group C (6 point scale, negative-positive order)

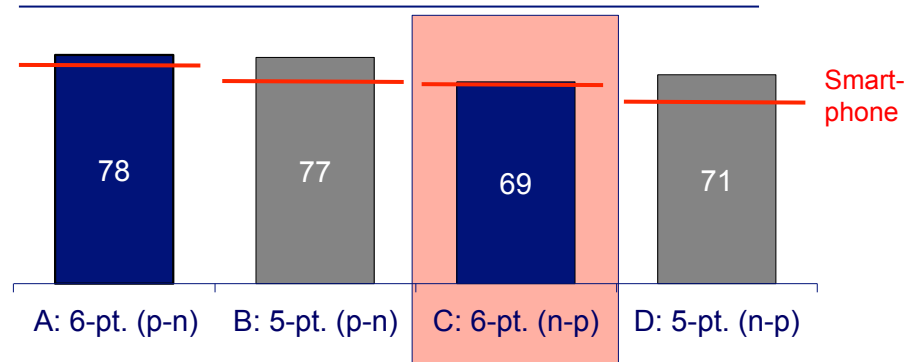
Study # 1

Degree of satisfaction with the interview (standardized values, 0-100)¹⁾

Topic of the interview

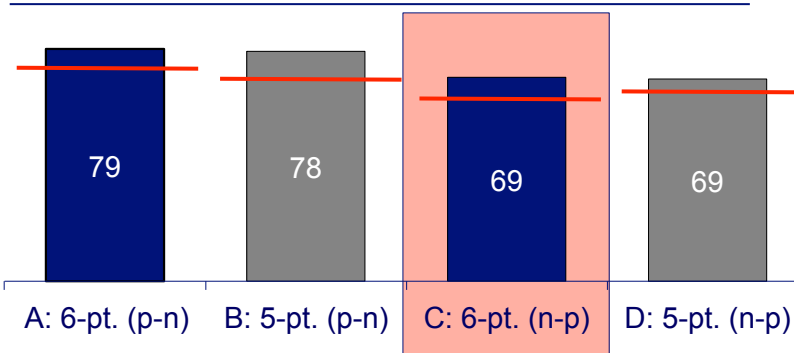


Length of the interview

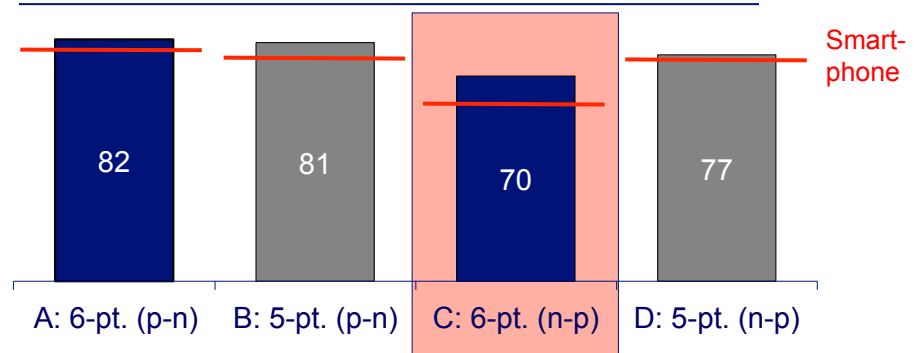


Range:
9-13

Graphical elements (images)



Interview overall



1) How satisfied are you with the following characteristics of the questionnaire?

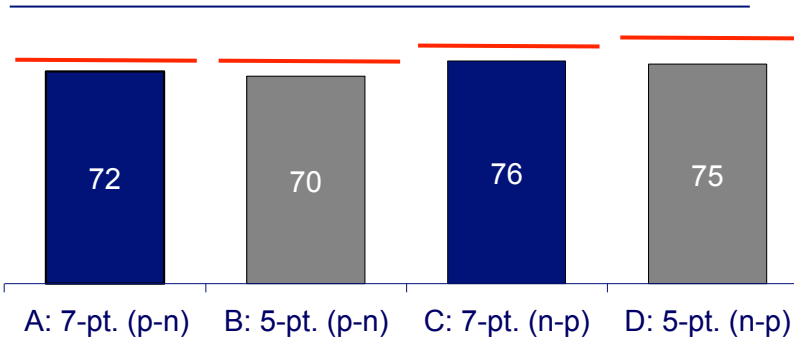
Source: **exeo** Strategic Consulting AG / Rogator AG

Study #2 does not confirm the findings from study #1: the range of results is more narrow, the range of results across test groups is smaller

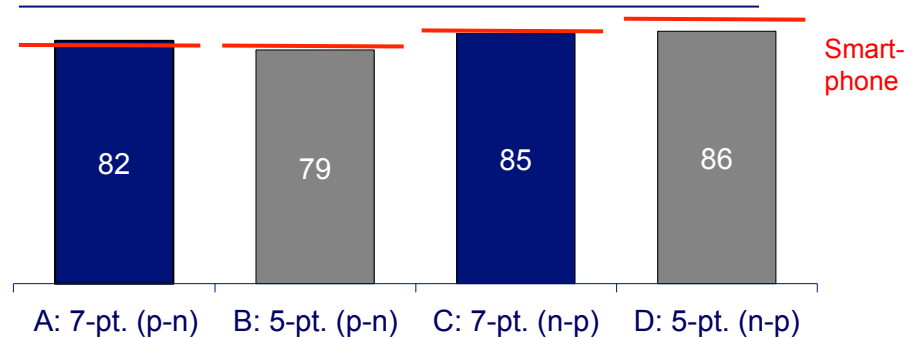
Study # 2

Degree of satisfaction with the interview (standardized values, 0-100)¹⁾

Topic of the interview



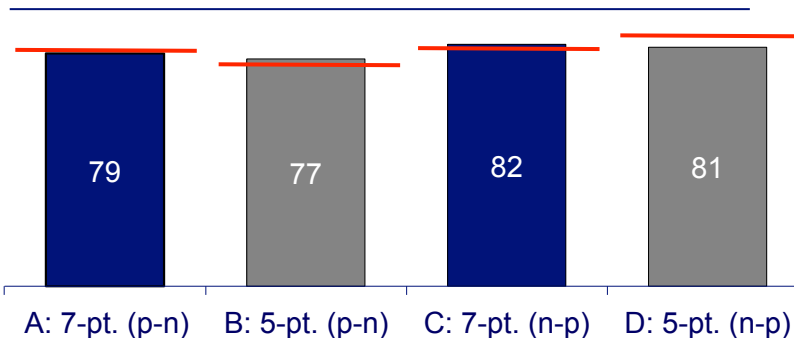
Length of the interview



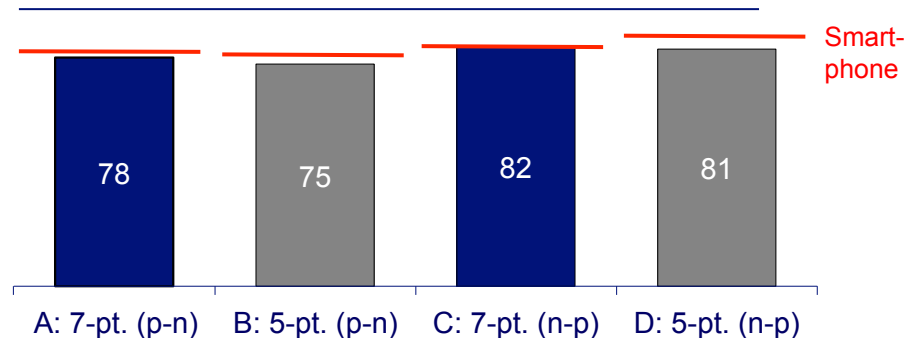
Smart-
phone

Range:
4-7

Graphical elements (images)



Interview overall



Smart-
phone

1) How satisfied are you with the following characteristics of the questionnaire?

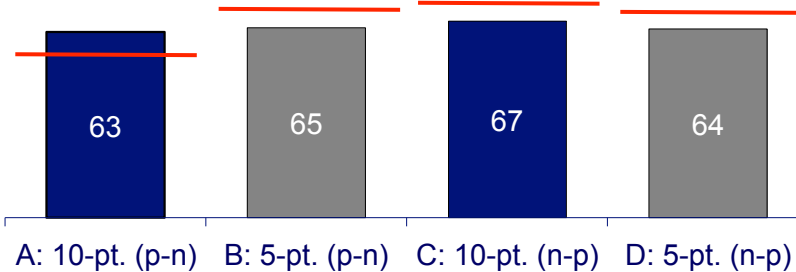
Source: **exeo** Strategic Consulting AG / Rogator AG

Study #3: Variance between treatment groups strongly reduced

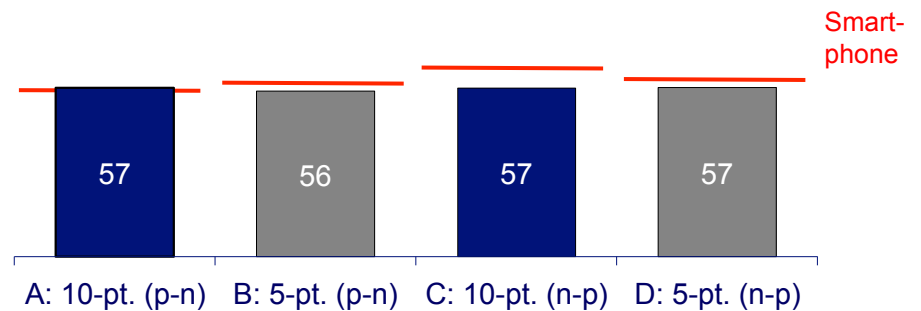
Study # 3

Degree of satisfaction with mobility options (standardized values, 0-100)¹⁾

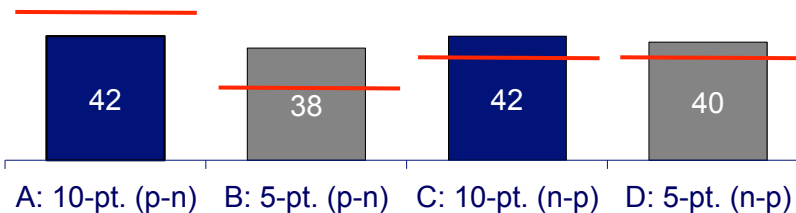
... supply of journeys with long distance buses



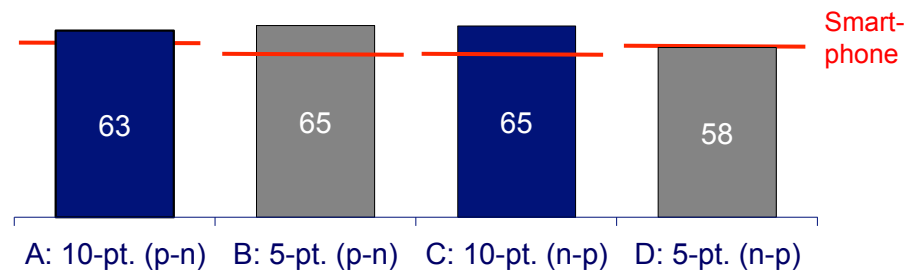
... market share (90%) of company Flixbus



... motorway toll for passenger cars



... motorway toll for trucks (in use since 2005)



1) How satisfied are you with the following characteristics of the questionnaire?

Source: **exeo** Strategic Consulting AG / Rogator AG

The device used to answer the questionnaire is strongly correlated with the age of the respondents

Cross tabulation device used for the survey vs. age group¹⁾

Study # 1						Study # 3					
Device used	< 30 years	30 -59 years	60+ years	Total*	Time (Min.)	Device used	< 30 years	30 -59 years	60+ years	Total**	Time (Min.)
Desktop	30 %	63 %	90 %	64 %	Ø 13,5	Desktop	61 %	72 %	93 %	75%	Ø 15,9
Tablet	8 %	10 %	4 %	8 %	Ø 14,4	Tablet	5 %	8 %	4 %	6 %	Ø 15,6
Smart-phone	62 %	27 %	6 %	28 %	Ø 15,7	Smart-phone	34 %	20 %	3 %	19 %	Ø 15,8
Total	100 %	100 %	100 %	100 %	Ø 14,2	Total	100 %	100 %	100 %	100 %	Ø 15,8

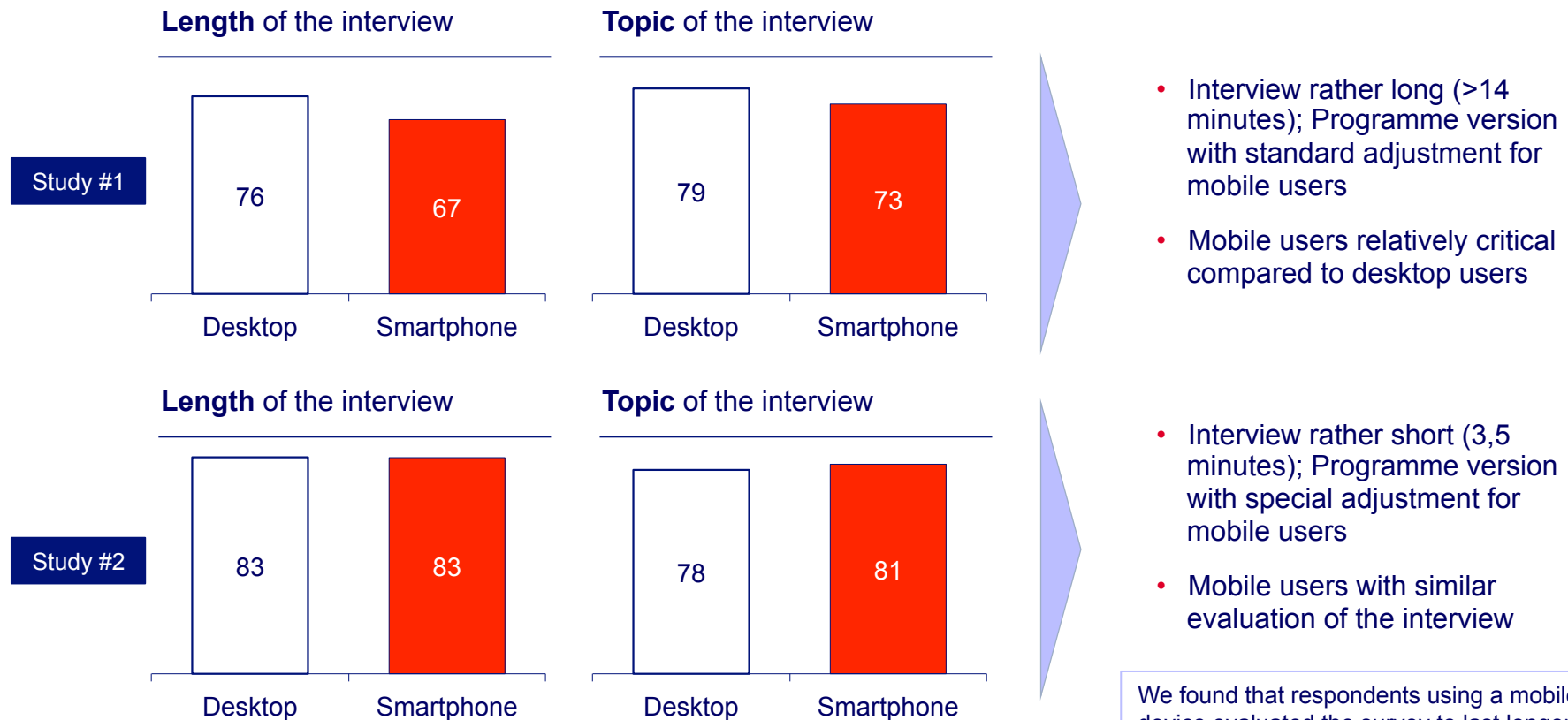
* Chi-Square: $p < 0.001$;

** Chi-Square: $p < 0.001$;

1) How old are you?

For longer interviews (study #1) respondents via smartphone are more critical concerning the length of the interview and the overall evaluation

Degree of satisfaction with the interview (standardized values, mean, 0-100)¹⁾



We found that respondents using a mobile device evaluated the survey to last longer than respondents using a computer: de Bruijne, & Wijnant (2013).

1) How satisfied are you with the following characteristics of the questionnaire?

Source: **exeo** Strategic Consulting AG / Rogator AG

Agenda

1. Introduction: the rise of mobile
 2. Research Method: online testing
 3. Results: three different studies
 - 4. Discussion: risk when changing scales**
 5. Outlook: what are the learnings?
-

Sometimes difficult: The step from a desktop design to a mobile design

More devices = more complexity in programming

Screenshot desktop

We would like to ask for one more minute of your valuable time to evaluate the interview.

How satisfied are you with the following characteristics of the questionnaire:

	1= very unsatisfied	2	3	4	5= very satisfied
Topic of the interview	☐	☐	☐	☐	☐
Length of the interview	☐	☐	☐	☐	☐
Graphical elements (images)	☐	☐	☐	☐	☐
Clarity of the questions	☐	☐	☐	☐	☐
Interview overall	☐	☐	☐	☐	☐

Continue

Screenshot Smartphone

How satisfied are you with the following characteristics of the questionnaire:

	1= very satisfied	2	3	4	5= very dissatisfied
Topic of the interview	☐	☐	☐	☐	☐
Length of the interview	☐	☐	☐	☐	☒
Graphical elements (images)	☐	☐	☐	☐	☐
Clarity of the questions	☐	☐	☐	☐	☐
Interview overall	☐	☐	☐	☐	☐

Continue

How satisfied are you with the following characteristics of the questionnaire:

	7= very dissatisfied	6	5	4	3	2	1= ver satisfie
Topic of the interview	☐	☐	☐	☐	☐	☐	☐
Length of the interview	☐	☐	☐	☐	☐	☐	☐
Graphical elements (images)	☐	☐	☐	☐	☐	☐	☐
Clarity of the questions	☐	☐	☐	☐	☐	☐	☐
Interview overall	☐	☐	☐	☐	☐	☐	☐

Continue

Note: Scale is not equidistant, the right scale of the scale is not shown in full

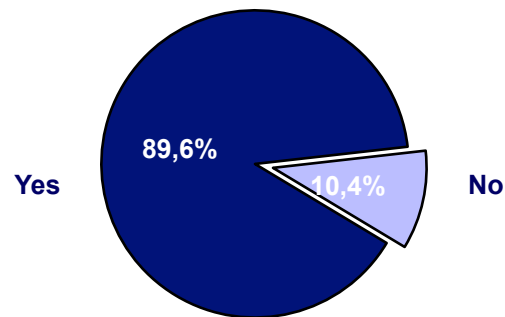
Almost one third of the respondents using a smartphone indicate discomfort when reading the scale

Study # 2

Share of respondents with discomfort when reading the scale

Answered via Desktop

Could you see the full range of the scale in the previous question without screen-scrolling?

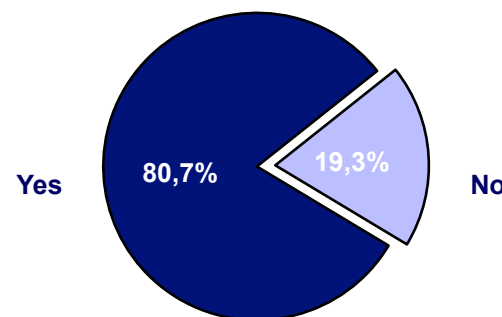


7 pts. scale (10 %) vs. 5 pts. scale (11 %)

Share of total sample: 58 %

Answered via Tablet

Could you see the full range of the scale in the previous question without screen-scrolling?

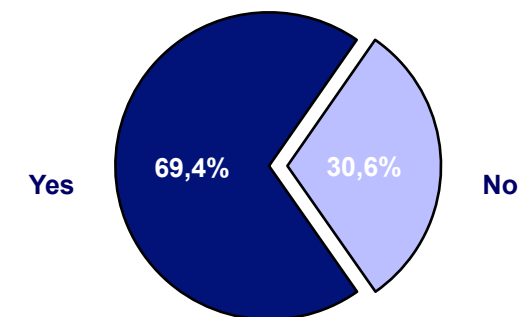


7 pts. scale (19 %) vs. 5 pts. scale (20 %)

Share of total sample: 18 %

Answered via Smartphone

Could you see the full range of the scale in the previous question without screen-scrolling?



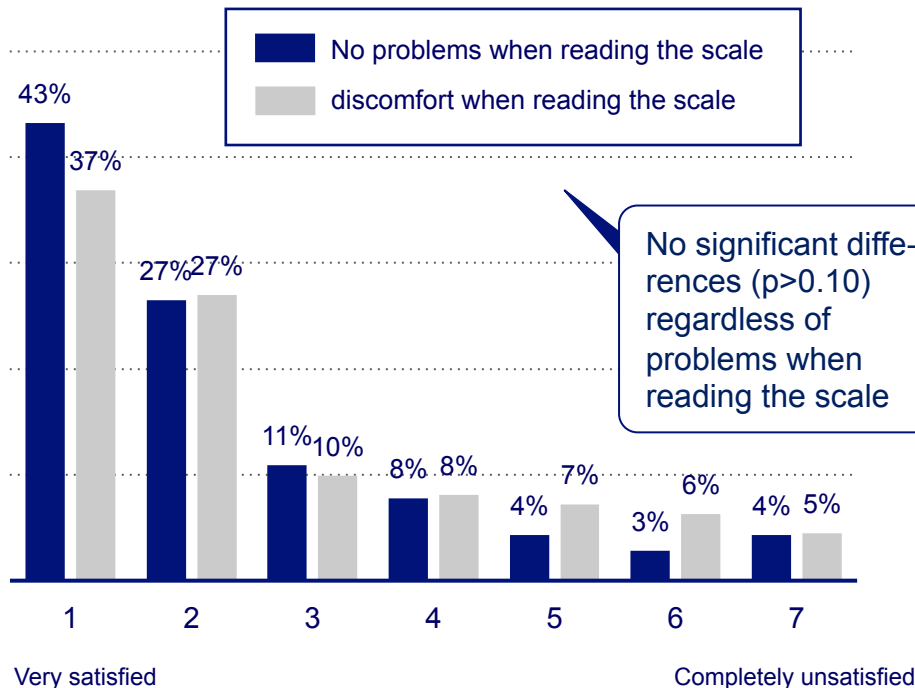
7 pts. scale (30 %) vs. 5 pts. scale (31 %)

Share of total sample: 24 %

In case of a 7-point scale (negative – positive) discomfort when reading the scale leads to changed distribution of answers

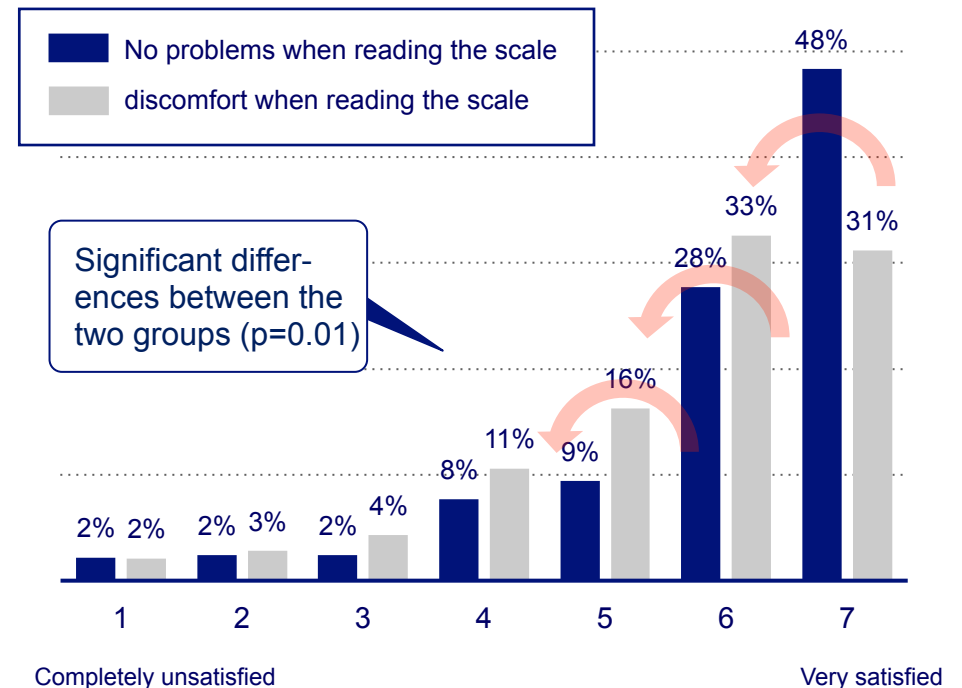
Degree of satisfaction with the interview (Distribution of rating results)¹⁾

A: 7-pt. scale (positive - negative)



Top-2: 70 % vs. 65 %

C: 7-pt. scale (negative - positive)



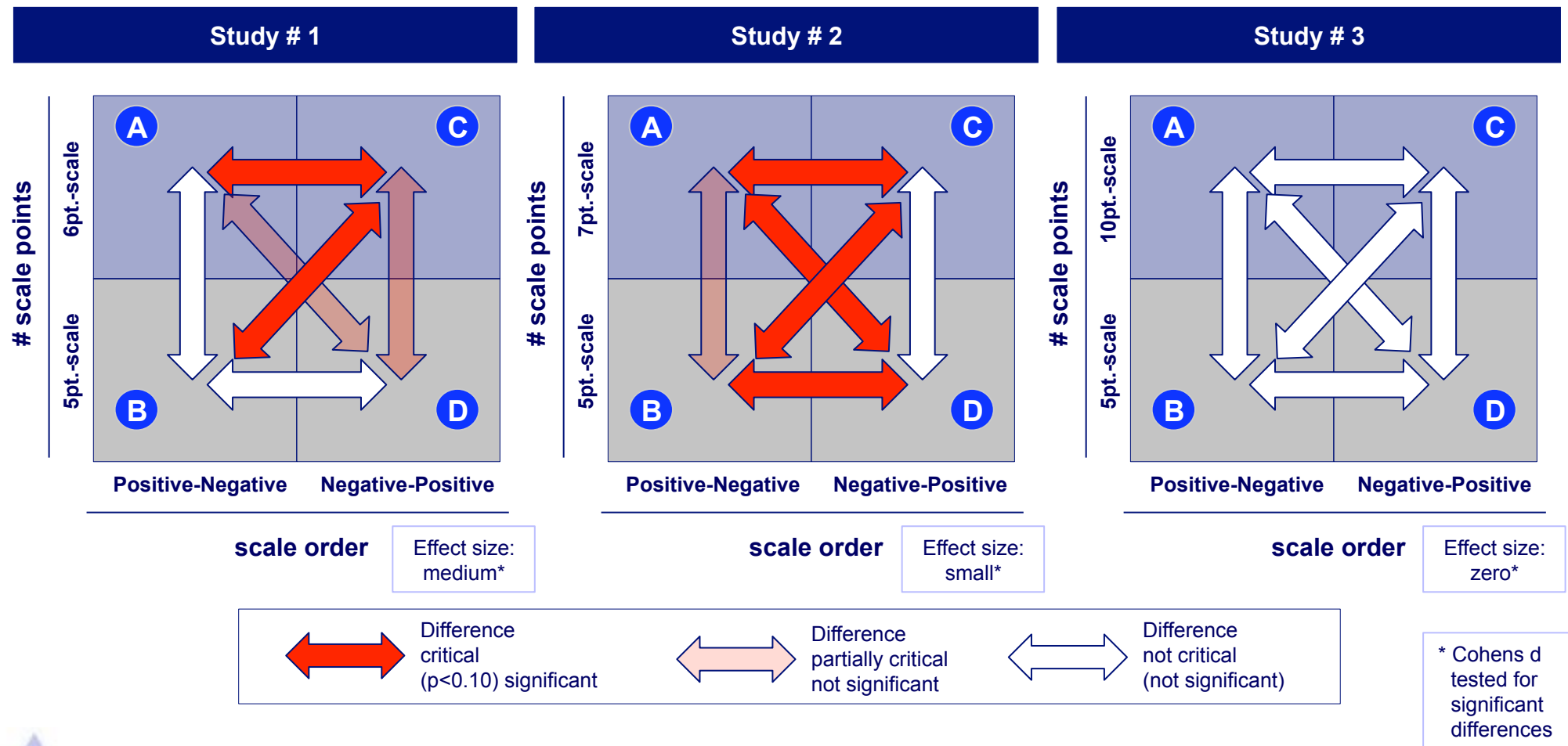
Top-2: 75 % vs. 64 %

Agenda

1. Introduction: the rise of mobile
2. Research method: online testing
3. Results: three different studies
4. Discussion: risk when changing scales
5. Outlook: what are the learnings?

Results of scale tests (2*2) are not robust: overall, it is difficult to generalize effects

Rating scale results (degree of satisfaction) across 3 different online studies



From a problem oriented perspective towards a solution-oriented perspective

GRIT REPORT, QIII, 2016

	Rank	Labels	In Use	Under Consideration	Total Interest
Mainstream	1	Mobile Surveys	75%	16%	91%
	2	Online Communities	59%	23%	82%
Wide Adoption	3	Social Media Analytics	52%	24%	76%
	4	Text Analytics	46%	30%	76%
	5	Webcam-Based Interviews	43%	22%	65%
	6	Mobile Qualitative	42%	26%	68%
	7	Big Data Analytics	38%	31%	69%
	8	Micro-surveys	35%	25%	60%
	9	Eye Tracking	35%	21%	56%
	10	Mobile Ethnography	33%	27%	60%
	11	Behavioral Economics Models	29%	25%	54%
Niche	12	Research Gamification	25%	29%	53%
	13	Prediction Markets	24%	23%	47%
	14	Facial analysis	24%	21%	45%
	15	Crowdsourcing	16%	21%	37%
	16	Neuromarketing	16%	19%	35%
	17	Virtual Environments/Virtual Reality	14%	24%	38%
	18	Internet of Things	14%	26%	39%
	19	Biometric Response	12%	19%	31%
	20	Sensor/Usage/Telemetry Data	11%	19%	31%
	21	Wearables Based Research	10%	27%	37%

„We feel the same pain that many clients and suppliers do in trying to migrate to new modes or incorporate emerging best practices in mobile-friendly designs“

Conclusion 1

- Across test groups top-2-values differ up to 22 Percent points (5pt/6pt) – caution when reporting top-2 values!

Conclusion 2

- Although standardized, results from different scales might be not compatible: do change scale format only, if necessary!

Conclusion 3

- Moving from online (desktop) to mobile sample increases the risk of discomfort when reading scales

Conclusion 4

- Problems when interviewees read scale can affect results in particular when scales are bigger and have neg.-pos. order

Conclusion 5

- Challenge: determine a scale design that fits for both groups, mobile and traditional online respondents

Literature

- Buskirk, T. D. (2015). Going Mobile with Survey Research: Design, Data Collection, Sampling and Recruitment Considerations for Smartphone and Tablet Based Surveys. Shortcourse presented at the Journal of Official Statistics Anniversary Conference, 2015. Stockholm, Sweden. Available at: http://www.scb.se/Grupp/Produkter_Tjanster/Kurser/_Dokument/JOS-2015/buskirk-FINAL-participant-JOS2015ShortCourseBus-kirkJUNE2015.pdf
- Buskirk, T. D., & Andres, C. (2013). Smart surveys for smart phones: Exploring various approaches for conducting online mobile surveys via smartphones. Survey Practice, 5(1).
- Christian, L. M., Parsons, N. L., & Dillman, D. A. (2009). Designing scalar questions for web surveys. Sociological Methods & Research, 37(3), 393-425.
- Dawes, J. G. (2012). Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. International Journal of Market Research, 50(1), 61-77.
- de Bruijne, M., & Wijnant, A. (2013). Comparing survey results obtained via mobile devices and computers: An experiment with a mobile web survey on a heterogeneous group of mobile devices versus a computer-assisted web survey. Social Science Computer Review, 31(4), 482-504.
- Garland, R. (1991). The mid-point on a rating scale: Is it desirable. Marketing bulletin, 2(1), 66-70.

Literature

- GRIT Report (2016). Greenbook Research Industry Trends. 2016 Q3–Q4 Tetrieved from <https://www.greenbook.org/images/GRIT/2017/GRIT2016-Q3-4.pdf> (accessed Feb 23, 2017)
- Leung, S. O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales. *Journal of Social Service Research*, 37(4), 412-421.
- Maloshonok, N., & Terentev, E. (2016). The impact of visual design and response formats on data quality in a web survey of MOOC students. *Computers in Human Behavior*, 62, 506-515.
- Peytchev, A., & Hill, C. A. (2010). Experiments in mobile web survey design: Similarities to other modes and unique considerations. *Social Science Computer Review*, 28, 319–335.
- Rammstedt, B., & Krebs, D. (2007). Does response scale format affect the answering of personality scales? Assessing the Big Five dimensions of personality with different response scales in a dependent sample. *European Journal of Psychological Assessment*, 23(1), 32-38.
- Schwarz, N., & Hippler, H. J. (1994). The numeric values of rating scales: A comparison of their impact in mail surveys and telephone interviews. *ZUMA Arbeitsbericht*, Mannheim August 1994.
- Toepoel, V., & Lugtig, P. (2015). Online surveys are mixed-device surveys. Issues associated with the use of different (mobile) devices in web surveys. *Methods, Data, Analyses*, 9(2), 155-162.