dr Marek Muszyński
Polska Akademia Nauk

Survey fatigue and data quality: reliability, validity, careless responding, and response styles.

The effects of questionnaire length on self-report data quality were studied at least from the seminal article of Herzog and Bachman (1981). The common-sense knowledge is that length should affect the last items responded to a greater degree (Jabine et al., 1984), which is exactly what was found by Berry et al. (1992) and Galesic & Bosnjak (2009) who found that participants responded faster and more uniformly (straightlining) a the end of a survey, and Krosnick et al. (2002) found that at the end "don't know" responses are more frequent.

However, despite the long history of this line of research the research problem is still far from being solved. Experimental studies are rare in this topic and even if they are conducted there are still some important research lacunas.

All main limitations can be addressed by analysing PISA 2012 data as the PISA 2012 rotational questionnaire design constitutes a natural experiment situation and is quite unique among the PISA editions as it was not used in PISA 2009 and before (OECD, 2010) and it was changed to a rotational assignment of non-overlapping blocks starting from the PISA 2015 cycle (OECD, 2015). This situation is beneficial for research practices as most of the survey fatigue articles confounded content with position and the PISA 2012 rotational design enables to overcome this limitation.

The idea is based on an assumption that students that received certain items later on in the survey would be more fatigued, hence would yield more careless/stylistic responses resulting in data of lower quality (similar to the study by Borgonovi & Biecek, 2016).Our study will also supply previous studies by presenting a wide range of careless/insufficient effort responding (C/IER) and response styles indices as measures of fatigue effects (Bowling et al., 2020). Furthermore, we will provide an in depth analysis of the relation between fatigue effect and scales' reliability and validity.