Understanding response styles in self-report data: consequences, remedies and sources.

We all know personality tests or surveys, where we are asked about our interests or political attitudes. Most often these questions are asked in an easy to answer format, where we are asked how much we like (or dislike) certain thing or how much we agree (or disagree) with a given statement. The undeniable appeal of this format of questions is most certainly shared by the researchers around the world. The so-called rating scales (which is a more formal name to those agree-disagree questions) are used across many scientific disciplines, not only in social sciences, but also in medicine and engineering. Moreover, these methods are not only cost-effective and very flexible, but also yield valid results, furthering our understanding of the social world (and not only).

However, despite the huge advantages this method is not free from many threats that can undermine the validity of the gathered results. Among many such threats one of the most serious is when respondents do not answer directly to the questions they are asked but do something different, at least through a part of the response process. This "something different" can range from answering totally at random, selecting only one answer for each of the questions or endorsing only a very limited number of responses, regarding the questions asked. So, someone can just cross all the options "agree" for every question and finish the survey or chose only between "agree" and "strongly agree" options. This kind of behaviour is formally called "response style". It is defined as a tendency to overuse a given response option regardless the question content. This behaviour is believed to be a very serious threat to the validity of the results obtained from numerous scientific studies. In spite of the popularity of self-report data the exact consequences of response styles presence is not known. Furthermore, little is known how to measure the response style contamination in our data and control for its detrimental influence. Finally, we are not so sure what are the sources of response styles in data, namely what causes people to respond in this way.

In order to answer the above questions we have conceived this project. Under the auspices of this grant we aim to search for the consequences, remedies and sources of response styles, amassing a sizeable amount of knowledge to the field. To this aim we will use a wide plethora of scientific methods. First of all, we will use simulation studies, that is to say we will simulate different data conditions, e.g. different level of different response styles, and check what are the consequences of such data contamination for the conclusions that we draw form the studies' results. Moreover, we will look for statistical remedies to account for the response styles in the data. Many methods were proposed to date but no systematic comparison of the remedies has been conducted as yet. To find the best statistical remedies how to measure and control for response styles we will again use simulation studies as well as we will analyse the data stemming from large international projects (so-called secondary data). What is more, we also plan to generate our own data in a series of experimental studies. We will manipulate the scale characteristics like e.g. number of response categories, as well as task difficulty (both objective and subjective), motivation to participate, topic involvement and attentiveness among many others. We will also measure respondents' cognitive abilities. Everything in order to find the sources of the response styles.

We hope that our project will greatly enhance our understanding of the response styles and that we will be able to control for them statistically in substantive analyses.