



研究与发展部
Research and Development Division
Professional DynaMetric Programs, Inc.

数据质量管控和可信度研究方法流程

Data Quality Control and Integrity Methodological Process

作者 / By:
Lauren Matheny, PhD, MPH.
2022诺夫梅伯 / November 2022

数据质量管控和可信度研究方法流程

在以电子方式收集数据时，尤其是收集大量数据时，经常会遇到关于某些问卷回应者答案的数据质量和可信度问题。对于任何类型的数据收集，通常都会遇到以下几种状况，有些问卷回应者可能会匆忙完成问卷中的问题，而并没有太注意他们的答案（超速者），在完成问卷答题时变得分心（分心的受访者），或者在问卷答题时，出现可能系统化或可识别模式化的不合理答案（直线化/回应可信度的模式化）。尽管这些问题在数据收集中很常见，但由于多种因素，包括能够识别出其中一些问题的复杂性，以及全面识别所有潜在问题所需要花费的成本和大量时间，这些问题往往被忽视了。

在对ProScan问卷工具的评估和质量的持续投资过程中，根据多重标准对超过 29,000人提供的问卷回应的数据可信度问题进行评估，首先针对“超速者”问题对所有数据都进行评估，对于每份问卷回应，清楚记录问卷完成时间，并评估短于先前已经定义的的阈值时间。此外，还评估每份问卷回应完成时间是否异常长，这被称为“分心的反应者”。如果一份问卷回应时间超出了先前定义的可接受阈值，该问卷则会用一个使用指标变量进行标记。

所有的问卷回应都以33种不同的系统化回应模式进行了额外的评估，例如，问卷中每个题都选择相同答案的问卷回应者，这种被称为“直线型”，会被标记为“模式化回应可信度”问题。同时，对所有数据也进行无意义回应（回应可信度的不一致性）评估，例如，如果在一个活动水平量表中，其中一个题目询问问卷回应者是否能够“毫无困难”地走上4层楼梯，那么，假如他同时也回答，走上1层楼梯是“极度困难”，这样的答案就是不合理，这个无意义回应的相同逻辑应用于ProScan问卷中16个题目。

总体而言，ProScan问卷工具根据 48 项标准对问卷回应可信度和质量管控进行评估。虽然这种全面性的质量管控不是数据收集的一般现状，但重要的是，ProScan问卷的每一份回应答案都有进行这些可信度检测，在一份问卷回应上标记出这类的问题启动了质量评估的下一步，这代表了更进一步的调查或重新填写问卷的需要性。执行这些严格的评估有助于确保维持ProScan问卷工具在个性特质评估上是可信赖和可靠的来源。