**预测试**in目标人群→删除或修订用词不当的条目

真实场景视频库(Shen et al., 2020)——自动驾驶中的经典情景（视频）

n=102-1-1

**·初步项目分析(Item Analysis)**

·传统项目分析(Traditional Item Analysis/Classical Test Theory，CTT)方法

·项目反应分析(Item Response Analysis/Item Response Theory，IRT)方法

→删减条目13、23、39、40、41

**·探索性因子分析(Exploratory Factor Analysis)**

·主成分分析(Principal Component Analysis，PCA)方法

·Promax 斜交旋转后得到因子负荷矩阵

→删改后保留38个条目

→量表调整为6个因子，并对条目进行了重新分配

**·访谈**

→根据用户的理解偏差与反馈，对部分条目做出语言上的修改

（文中未详细介绍，不知道是否要展开）

**初步项目分析(Item Analysis)**使用传统项目分析(Traditional Item Analysis/Classical Test Theory，CTT)+项目反应分析(Item Response Analysis/Item Response Theory，IRT)

用两种优势互补的项目分析方法，对量表条目难度、区分度、参测参数等进行分析，删减那些表现不佳的条目。例如，如果一个条目的难度指数非常低，表明几乎所有受测者都能答对，那么这个条目可能无法提供有用的信息，因此可能被删减。

其中，CTT：

**难度指数(facility index)：**对于非对即错的题目而言，是选对答案的人数／全部答题的人数。最后的答案在0到1之间，数字越高，代表题目越简单，数字越低，代表题目越难。如果一个条目的难度指数太高或太低（通常认为大于0.9或小于0.1为不合适），它可能无法有效区分不同能力的受测者。

**区分度(discrimination index)**：是用来区分同一个题目对于成绩好的受试者和成绩差的受试者之间水平差异的指数，做区分度的时候，需要把受试群体按照成绩高低分成地低分组，中间组和高分组。区分度的数值从-1到1之间浮动，-1表示所有低分群体都选择了正确答案而所有高分群体都选择了错误答案，0表示高分组和低分组之间没有差异，0.3的区分度一般是比较可接受的。如果两道题目的难度相同，区分度高的题目往往质量更高。

**点双列相关：**计算每个条目与总分之间的相关性，以评估条目对总分的贡献。

IRT：

**难度参数（b）**：评估受测者在某个特定能力水平下选择某个答案类别的概率。

**区分度参数（a）**：衡量条目对受测者能力的敏感度。高区分度的条目能够更精确地测量受测者的能力差异。

**猜测参数（c）**：评估受测者随机选择某个答案类别的概率。

采用**主成分分析(Principal Component Analysis，PCA)**方法完成**探索性因子分析**

降低数据的维度，同时尽可能保留原始数据的变异信息

数据清洗与标准化→计算协方差矩阵或相关矩阵→计算特征值和特征向量→选择主成分并解释→因子旋转→构建因子得分

Promax斜交旋转后得到的因子负荷矩阵用于删减条目，解决交叉载荷问题，以降维实现优化量表结构