

关于标准得分

新“日本語能力测试”（新考试）中引进的“标准得分”，是为了更加公平地评价考生的日语能力和测试结果（得分）而采用通用标准表示的得分。利用标准得分，无论考生参加的是哪一次考试都可以保证在同一标准下测定其日语能力。以下对标准得分的好处及其计算过程做一下说明。

【标准得分的好处】

即使在出题过程中经过了专家的严密分析与研究，也很难保证每次试题的难易度完全一致。为此，如果单纯地计算考生正确答题所得分数（纯分数），考试得分便会依存于考试的难易度。这样参加考试的时间不同或者试题的难易度不同时，即使考生拥有同样的日语能力，考试所得分数也不会相同。因此，我们必须考虑到用纯分数表示考生日语能力时，由于受试题难易度的影响，纯分数有可能反映不出考生的实际日语能力。

与此相比，新考试中引进的“标准得分”在能力测定方法上有其非常好的特点，即“不依存于考试的难易度。独立地对日语能力进行评价，可根据统一的标准予以数值化。”根据这一特点，只要考生的日语能力相同，不论何时参加考试，都可以得到同样的分数。引进标准得分后，只要是相同级别的相同得分分类，就可以把两次不同考试中算出的“标准得分之差”作为“日语能力之差”，独立于考试的难易度而做出解释。

【标准得分的计算过程】

不依存于考试难易度的考试得分（标准得分）的具体计算程序，是以称为项目反应理论(Item Response Theory; IRT)的统计学考试理论为依据的。它完全不同于根据正确答题数量计算出分数（纯分数）的方法。

标准得分是由答题情况（答题正确还是错误）、也就是“解答类型”得出数理性结论的。例如，由10道（项目）试题组成的考试中，最多有1024种（ 2^{10} 种）解答类型。在新考试的标准得分计算过程中，基于这些解答类型，把考生的日语能力按每个得分分类划归0~60分（N4、N5的“语言知识<文字、词汇、语法>阅读”中为0~120分）的标准。由于把10道试题会有的多达1024种的解答类型（实际考试中会有更多）分成了61组，不同的解答类型有时会得到相同的标准得分。因此，即使两位考生的正确答题数量及解答类型有所不同，其标准得分有时也会相同。相反，即使正确答题数量相同，其标准得分有时也会因解答类型不同而有所不同。

关于标准得分计算过程的详细情况，请参照下列文献。

【文献】

池田央（1994）《现代考试理论》 朝仓书店

丰田秀树（2002）《项目反应理论“入门篇”——考试与测定的科学》朝仓书店

渡边直登、野口裕之（主编）（1999）《组织心理测定论——项目反应理论的前沿》 白桃书房

关于合格分的设定

新考试中，在设定合格与否的判定标准时，基于统计分析结果，使其与旧考试的合格判定标准有了一定的整合性。由此，日语能力能够通过旧考试中的1级、2级、3级、4级的考生，分别可以视为其日语能力能够通过新考试中的N1、N2、N4、N5。对于旧考试中不存在相应级别的N3，通过对旧考试中2级和3级的日语能力水平的统计学分析，设定了N3的合格分，使其介于旧考试的2级合格分与3级合格分之间。