

国际测评为教育科学评价带来什么启示

去年10月13日，党中央国务院发布新时代改革教育评价的总方案。《总方案》明确了教育评价改革的指导思想和改革的原则要求。完善立德树人体制机制，扭转不科学的教育评价导向，坚决克服唯分数、唯升学、唯文凭、唯论文、唯帽子的顽疾，提高教育治理能力和水平，加快推进教育现代化、建设教育强国、办好人民满意的教育。

如何建立科学的评价体系，这一直是学界关注的话题。上海自从首次参加经济合作与发展组织(OECD)的PISA测试至今，已经多次参加包括PISA、TALIS、TALIS Video Study等在内的多种教育国际测评。

这些测评让我们看到了中国上海基础教育的发展成就、上海教师的贡献和教育体制的优势，增强了我们的自信。也吸引了世界的目光，许多发达国家派出大批教育行政官员、专家和教师来上海学习探秘，还邀请上海教师到英国去示范教学、合作交流。

然而，从国际测评中我们获得的启示更加丰富。我们不仅看到了上海教育的长板和短板，还学会如何科学设计测评方案，并且通过测评大数据的使用，来为教育发展的施策提供更多参考依据。

■张民选

上海已经参加过多次PISA测试，这些测评让我们看到，我们的学生在阅读、数学和科学等基础素养(literacy)方面非常优秀，甚至达到全球第一。但是，测评也显示出我们的学生在“可迁移的跨学科素养(胜任力)”方面成绩不够好，折射出我们对这一素养既不善教，也不会评。

从2003年起，经合组织在PISA测试中，增设了“21世纪可迁移的跨学科素养(胜任力)”测试。首先设立了“问题解决能力”测评；2012年增设了“在计算机上解决问题”的测试和“财经素养”测试；2015年测评学生的“协作性问题解决”的能力；2018年测评学生的“全球胜任力”；2021年还将测评学生的创新能力。这都对我们的教学和评价提出巨大的挑战。

比如，众所周知“解决问题”的能力重要，但是测评反映出我们的老师会教语文、数学、外语，却不会教学生解

信息技术、大数据和人工智能的发展，为我们的教育测评增加了更多的维度。以PISA测试为例，80个国家60万左右15岁的孩子参加测评后形成的大数据，包括了学科、学校、教师、国家以及学生家庭的情况，这些为我们提供了很多分析的资源。

分析这些大数据，我们可以发现教学中的短板。比如，我们的老师不善于使用信息技术让学生自主学习。调查发现，44个国家中有52.7%左右的老师每周都用信息技术让学生完成作业，而我们的这一比例只有24.3%。

同时，TALIS的视频研究，让8个国家的近千名数学教师“同课异

各类国际测评的数据让我们深刻认识到，教育是一个复杂的系统，要通过综合评价和过程评价，改变教育过程的“黑箱”效应。

现在的教育仿佛“黑箱”，大家只看到一个小孩进去再出来。但是在这个过程中究竟发生了什么？我们往往过于关注教育的输入，也关注这个“黑箱”中出来的结果，但是“黑箱”中究竟哪些因素、究竟如何影响教育，却缺乏研究。

一些国际测评也给了我们启示。比如，PISA测评中的阅读、数学和科学测试也许能够表现一个国家或者地

评优性评价需要减少，更应倡导政策改进性评价

更科学的测评应该是减少评优性、终结性的评估，增加过程性评估。

根据测评结果对“黑箱”中的因素分析，我们可以发现，对学校频繁的评优性评估并不能提高教育质量，反而会导导致学校的生源竞争，鼓励学校加强标准化测试和外部测试的管控。

芬兰也曾遭遇过频繁的评优性评估带来的教育不均衡以及教育质量下降的问题，但是从1990年代起，他们树立了“少考试、多学习”的理念，加强校内老师的课堂检测，并且增加学生的反馈，来提升教育质量。同时，学生评价的重点也不再是考试与分数，而是描述性的学习过程的评价，包括学习行为和学习参与度。政府对教育质量的把控也不再是频繁的高利害考试和对学校教师的不断评优，而是通过抽样性外部评价来实现。芬兰政府每学期都抽样10%的学校和学生对各个方面进行评估，这既减轻学生负担，又能把控学校质量。

同时，分析“黑箱”中的因素，我们也发现，学校质量与投入有关，也与校长教师的努力创新有关。政府应根据

区的教育发展和学生学业的水平，而学生和校长的问卷却能解释其原因与路径。问卷涵盖了学生个人学习投入的时间和方式、学校内部的差异与学校内部的要素、家庭经济教育的影响与家庭迁徙的影响以及政府财政的投入程度与投入方式等。

世界银行组织的“SABER”(教育成就系统评估研究)给我们提供了

一个有益的框架。这个框架告诉我们，对学生、学校和教育的质的评价不要只看学生的成绩，还要考虑教育过程“黑箱”中的要素，如政策层面的教育制度、教育财政、学校职权、教师政策，教育机构层面的学校经费、学校课程、教师发展等，督导问责层面的政府职责、学校条件、教师效能和学生评价等。

这个框架告诉我们，对学生、学校和教育的质的评价不要只看学生的成绩，还要考虑教育过程“黑箱”中的要素，如政策层面的教育制度、教育财政、学校职权、教师政策，教育机构层面的学校经费、学校课程、教师发展等，督导问责层面的政府职责、学校条件、教师效能和学生评价等。

因此，要大力倡导科学的过程监测，通过计算机、大数据和人工智能的帮助，我们可以通过教育过程监测，及时记录教育政策、教育方法、教师教学和学生学习实施成效、特性与问题，调整政策、资源、方法，帮助教师因材施教，为学生精准推送学习资源。让基础教育更加公平优质发展、让学生全面而有个性发展。

(作者为教育部国际教育研究与咨询中心主任、上海师范大学国际与比较教育研究院院长、联合国教科文组织教师教育中心负责人)

国际测评让我们意识到教育中的不足,包括教育评价的落后

决问题，也很少有人学习和研究“解决问题”的过程、结构、技能、态度要素。

那什么是解决问题的能力？如何测评解决问题的能力？

专家发现：问题解决能力就是在没有明确的解决办法的前提下，个人有意愿投入认知过程，以认识、理解和解决问题情境和解决问题的能力。这种能力包含了人的意愿和认知的过程。

研究团队还发现，问题种类很多，但至少可分为“静态问题”和“动态问题”。所谓“静态问题”就是纸笔测试中，应有条件、必要条件都已具备，并且没有虚假条件的问题，只要学生选择正确有用的方法、并作某些计算和选择，就可以解答。

技术+大数据,为教育评价提供更多可参考的维度

构”，来分析教学中的长短。同一节“一元二次方程”课，同样用“课堂管理”“社会情感支持”和“教学质量”三大维度去评价。我们可以看到，成绩较优秀的国家，老师都关注“课堂管理”，其中上海教师做得最好，按四级评分，90%以上的教师都集中于3-4区间。但我们在课堂教学中给予学生的“社会情感支持”方面仅处中等水平，同时教师个体间的差异也很大。而日本教师不仅课堂管理出色，对学

通过教育的综合评价和过程评价,消除“黑箱”效应

生的情感支持也比我们高很多。另外，我们在学生认知参与度和师生课堂互动对话等方面也还有巨大的改进空间。

国际测评还告诉我们，我们中小学教师

的学位低、男教师少、学习的自主性积极性低。经合组织成员国初中教师

硕士学位比例已经高达44.6%。芬兰更是达到92%，而上海只有12.6%。同时，发达国家的男教师占33%，我们只有27%。

再看教师的专业化要素，这主要包



■钱军辉

放性。比如治理河水污染，未必通过生物、化学的办法解决，也可以通过经济杠杆比如高额处罚来阻断污染源，这就是开放性，即很多问题并非单一学科来解决的。

我们对“问题解决”这一跨学科、可迁移素养的研究不足，也导致我们教学中这一素养的欠缺以及评价的缺失。

在2012年的PISA测试中，我们同样发现很多国家学生的数学、阅读、科学的成绩与解决问题的能力成正比，但是我们却恰恰相反，而且差距还比较大。

国际测评也让我们发现，虽然我们的学生成绩很好，但是学生的学业负担也很重。2012年上海15岁学生平均每周的作业时间是15个小时。作业效果的最佳临界点却是10小时，超过以后，再多的时间对成绩提升的作用也很有限，反而占据了学生发展个性、锻炼身体的时间和空间，甚至还会造成学生厌学。

括教师的知识技能、教师间互助协作和专业自主性，以及社会给予教师的专业发展机会和社会地位与尊重。我们教师的专业自主性、特别是教师自主学习的积极性比较低。

在四分等级的国际比较框架中，芬兰教师的平均值为2.9，上海教师的平均值仅为1.1。这提示我们，原因可能与当下的教师培训组织有关，我们的培训进修几乎都是由行政部门和学校规定。久而久之，教师就不再考虑自己应该学习什么，怎么学习了。而教师自主学习积极性低，还会影响教师对学生自主学习精神和习惯的培养。

对中小学来说，尤其要关注均衡投入以及单位投入产出的测评，同时进行分层分类的评估。只有这样才能提升所有校长和教师的积极性。

我们在2012年PISA测试中发现，芬兰的学校校间的差异(方差)仅为8%，我们的初中学校有29%，高中(包括职业学校)达到了58%，初中平均为47%。这就提醒我们，不解决学校之间的投入不平衡，家长就会想尽一切办法去择校。相反，在均衡的情况下，校长和教师的努力程度就会大幅度提高，而择校压力则会下降。

2018年又实施了TALIS Video Study(视频研究)，开展八国教师课

堂教学行为进行评价研究。

作为教育评价一类方法，国际教育测评始于1960年代。

1961年，国际教育成就评价协会(IEA)就着手组织开展了第一次“国际数学测评研究”(FIMS)，有12个国家参加。1964年，又开始了六个学科教育国际测评研究。1977年，实施第二次数学测试(SIMS)。

1995年国际教育成就评价学会组织实施了第三次数学与科学测试(TIMSS)，并改称为“国际数学与科学趋势研究”(TIMSS)。

2000年，经合组织举办PISA，每三年测试一次，测评的对象为义务教育末端的15岁在校学生。测试的学习领域为：阅读、数学和科学。以后又增加了其他学习领域的测试。如问题解决、协作解决问题、财经素养、全球胜任力，等等。目前已经有79个国家/地区的60多万名学生参与测评和问卷调查。

经合组织在PISA测试中发现，教师除了家庭以外，对15岁学生

学业影响最大的单一因素。于是，经合组织又于2008年开始组织“教师教学国际调查”(TALIS)，每五年一轮，对教师进行问卷调查，44个国家/地区的26万名教师参加。

2014年，世界银行设立了SABER(教育成就的系统评价研究)。

2018年又实施了TALIS Video Study(视频研究)，开展八国教师课

堂教学行为进行评价研究。

教师除了家庭以外，对15岁学生

自从负责本科生学术项目以来，时常会面对学生们的苦恼。他们或是因为生活、学习上遇到的困难，或是因为遭遇各种难以抉择的选择而苦恼，当然，也有不少学生入学以后不能很好地调整自己适应大学生活。大多数时候我们在面对学生的问题时，还会面对他们的家长。而此时往往会发现，这些生理年龄已经进入成年的大学生们，心理年龄并未成年，更多的时候是父母也并不愿意放手让他们走向成年。

很多时候，我想对父母和学生都说同样的一句话，那就是：当孩子逐渐迈向成年时，尤其是孩子进入大学以后，父母和孩子更应该向朋友过渡，在这样一个走向成熟的“冲刺期”，父母既没有能力也没有必要“管”着孩子，这样，只会适得其反。

进入大学的年纪，在中国正好是法律意义上成年的十八岁。但是大多数在这个年龄距离心理上的成年，可能还有一定的差距。这一点从生理学的角度有据可依：负责理性思维的前额叶(front lobe)是大脑发育最晚的部分，据说男性前额叶要到30岁左右才发育成熟，女性也要到25岁左右。

也就是说，大学生虽然身体已经发育成熟，但是他们的大脑尚未发育到能够很好地管理好自己的情绪，当然，面对各种诱惑，他们很难抵御，至少很可能缺少理性决策的能力等。

但是，年轻人情绪强于理性，情感胜于理智，并不完全是坏事。古往今来有所成就的人，多在年轻时就立下不够理性、不够理智的志向。年轻人需要梦想、野心和冲动，这样才有前进动力。年轻人也需要一点盲目自信，这样才能勇敢向前。

所谓心理上成年，不是理性压倒非理性，而是理性更好地为非理性服务。而且，理性的成长不只是生理现象，更重要的是让孩子在尝试和纠错中不断学习并接受磨练。而大学正提供了相对安全的尝试和纠错空间。

在大学四年中，这些年轻人有很多诱惑要抵御，有很多事关人生走向的重要选择要决定。因此不少父母在孩子走进大学校园后继续操心，甚至遥控和操办，这也似乎有科学依据。但现实是，大多数父母早已经力不从心。同时，绝大多数父母远在天边，并不能像中学时代那样管着孩子。孩子上大学，简直就像风筝断了线，只能咬咬牙由他们去，最多只能在电话和微信里反复叮嘱不要打游戏、好好学习等等，至于有没有用，只有天知道。

于是很多时候，父母只能从“全面监护”转为“重点干预”；把握重大方向，防止幼稚决策。

但是重点干预也有问题。首先，父母在知识和经验上有局限。绝大多数父母，指导高三学生选择大学和专业就已经很吃力，别说是大学生指点迷津。而且父母也是有缺陷的人类，容易过度自信，认识不到自己的局限性，这往往就会成为父母与孩子冲突的根源。

即使是能力十足的父母，到孩子进入大学阶段，也应该允许他们做自己的选择。大学生也许尚未真正成为心理上的成年人，但是大学阶段正是他们走向成年的冲刺阶段。只有放手让孩子拿主意，孩子才会学习拿主意，才会更快成长。如果坚持操办孩子的各种“大事”，那么孩子很可能成为“巨婴”，最后走上社会，却要吃大亏。

可以说，大学是最后一块比较安全的人生实习基地。在大学里做傻事的后果，只要不是触犯法律，远比在社会上做傻事的后果轻。即使是像选择专业和职业这样的重大决定，大学时期的选择其实也不致命。只要人有决心，永远都有第二次机会。比如错过读研的学生，在工作若干年后可以重返校园。误入研究生院的同学，如果发现自己不适合学术，同样也可以反悔，毕业后选择进入业界。

人生不是赛跑，比的不是速度，人生道路的曲折只会让人更加丰富。

当然，这并不意味着大学生可以“摆脱”父母的意见，自我放飞。我认为，父母的意见仍然宝贵，但是不论是父母还是学生都应该意识到，这些意见仅仅只能起到参考作用。学生进入校园后，父母要慢慢退回到孩子朋友的位置，不再是负责为孩子导航的“监护人”。只有这样，父母才是孩子需要倾诉时最先打电话找的人，孩子才会意识到，父母是最无私、最关心、和他们之间最知根知底的朋友。

从监护人与被监护人关系，转型到朋友关系，是迟早的事。等到父母年老体衰，父母和子女的关系还会再次转型。到那时，已经学会拿主意、已经久经沙场的儿女要为父母拿主意。角色的自觉转型，恰恰体现了家庭代际交替的智慧。

当然，大多数父母可能还没有做好思想准备，和进入大学阶段的孩子做朋友。这个时候，大学生要用实际行动说服父母。最简单的，就是校园里，控制住打游戏的冲动，管理好学习和课外活动的时间，处理好个人感情，谨慎用钱拒绝提前消费等。寒暑假回家的时候，要跟父母聊聊自己的专业，聊聊自己的规划和职业前景。当然，得事先做好功课，用自信 and 准备，赢得父母的信心。

和父母做成朋友，其实正是心理上成年的标志。转型不可能一夜之间完成，但是从大学入学起，这个转型必须开始。

(作者为上海交通大学教授)

放手让孩子在大学进行「人生实习」