



# 大数据背景下的心理学: 中国心理学的学科体系重构及特征

喻丰<sup>①②</sup>, 彭凯平<sup>②③\*</sup>, 郑先隽<sup>②</sup>

① 清华大学经济管理学院, 北京 100084;

② 清华大学心理学系, 北京 100084;

③ Department of Psychology, University of California, Berkeley CA94720, USA

\* 联系人, E-mail: pengkp@mail.tsinghua.edu.cn

2014-09-12 收稿, 2014-11-25 接受, 2015-01-22 网络版发表

清华大学-中国科学院学部科学与社会协同发展中心 2014 年度先导项目(2014A03)资助

**摘要** 实验方法使得心理学思想成为了心理学学科, 并以此建构起了西化的科学心理学体系。而大数据时代的来临极大地影响了心理学科的研究方式, 使得心理学面临又一次的变革。大数据方法的特性与中国文化和中国人的认识论特征不谋而合, 这给予了中国心理学摆脱西化影响, 并重构中国化心理学学科体系以机遇。我们认为, 中国心理学的学科体系可以遵循天、地、人相互关联的构建原则。从研究文化与精神的心理表征, 到社会生活情境的制约影响, 再到个体的知、行、情、理、德的身心体验, 真正达到孔德所描述的人类最后科学的境界, 即“具有社会特性的生物个体在文化情境中的行为”, 这就是我们提倡的中国的心理学。而在这种新的理论体系之下, 以大数据为根本研究范式的中国心理学将会超越传统心理学中样本—总体、个体—规律、情境—实验、数据—行为的二维争辩特性, 创建新的心理学研究范式。

## 关键词

中国心理学  
大数据  
学科体系  
文化  
认识论

2013年7月31日, 美国纽约州的米歇尔·卡塔拉诺(Michele Catalano)家里迎来了6名美国反恐警察, 调查他们是否是恐怖分子。而原因则只是因为她在谷歌(Google)上搜索高压锅的信息, 而她的丈夫则同时在谷歌上搜索背包的信息(<http://www.theatlanticwire.com/national/2013/08/government-knocking-doors-because-google-searches/67864/>)。美国警方能够如此判断是因为他们将高压锅、背包和2013年4月15日的波士顿爆炸案联系在一起(<http://www.cbsnews.com/feature/boston-bombings/>)。当然, 能够迫使美国警方如此行动的可能性在于他们获取了卡塔拉诺及其丈夫的搜索数据。我们不讨论这样一些类似事件的伦理问题和政治影响, 但仅就日常生活而言, 大数据已经不止停留在口号上, 它切实地影响了我们的生活。

2008年谷歌公司的数据工程师们通过谷歌保存

的搜索记录中的词条尝试以复杂数据为基础的社会计算模型, 并得到它们与美国疾控中心(U.S. Centers for Disease Control and Prevention)的流感数据的关系<sup>[1]</sup>。这一结果能够有效地预测之后的H1N1疾病的发生, 甚至预测得比疾控中心更快。这种流行病学的工作被数据工程师讨论并且完成得如此之好令人震惊。类似的研究也广受关注和讨论<sup>[2,3]</sup>。

大数据的研究已经在科学界崭露头角, 近年来 *Science*, *Nature* 杂志均出版专刊讨论大数据在科学中的运用及其潜在问题。以上面H1N1的例子来说, 流行病学可能因这种大数据方法改变其研究的思维方式和方式, 那么其他学科又将如何? 社会科学中以渗透数学的计算方法或者说主要以计算方式来研究已经有一段时间了<sup>[4]</sup>, 借助定量化的数学方法, 社会科学可能向纯自然科学靠拢<sup>[5]</sup>。那么在这样的背景下心理

**引用格式:** 喻丰, 彭凯平, 郑先隽. 大数据背景下的心理学: 中国心理学的学科体系重构及特征. 科学通报, 2015, 60: 520-533

Yu F, Peng K P, Zheng X J. Big data and psychology in China (in Chinese). Chin Sci Bull, 2015, 60: 520-533, doi: 10.1360/N972014-00977

学又将如何发展?

Twitter等社交网站是心理学中大数据搜集和研究的主要平台。通过对2008年8~12月Twitter中用户语言的数学分析,研究者们从中抽取出了紧张、抑郁、愤怒、活力、疲劳以及困扰6个维度,组成了一个情绪剖面的描述方法。而如果将这种情绪剖面系统与美国股市震荡、国际原油价格、类似美国总体选举这样受媒体关注的主要事件以及感恩节这样的流行文化事件进行对比,研究者们发现,社会、政治、文化和经济事件领域都高度强调不同的公共情绪的维度和组合,公共情绪的变化也能成为某些社会和经济指标的有效预测平台<sup>[6]</sup>。通过对全球Twitter数据中积极情绪和消极情绪的语义的分析,研究者们也发现,个体的每日心情变化是存在节律的。清晨时最为愉悦,接着慢慢变差,直到深夜,情绪再度变好,而周末的情绪则比周中要好<sup>[7]</sup>。这与用传统调查方法得出的结论基本一致<sup>[8]</sup>。同时,积极心理学家马丁·塞利格曼(Martin E. P. Seligman)等人<sup>[9]</sup>根据其幸福的PERMA理论,建构相应的词库,根据Twitter平台的用户自然语言和地点信息,画出了一张美国的幸福地图,并用以比较。通过对语言的数学分析,他们同样建构出了能够有效预测幸福的自然事件集合,这种通过自下而上数据驱动方式的预测源可能要比以理论驱动方式得到的预测源要更为有效<sup>[10]</sup>。

除了社交网站,图书、传统博客、短信以及手机App都有可能成为心理学大数据的来源。有研究者对谷歌图书中的电子数据中蕴含的文化趋势进行了分析,通过对占人类整个出版图书4%的5195769本电子图书的数学、语言学和文化分析,他们阐释了从1800年到2000年间英语文化中词典学、语法、集体记忆、科技使用、名利追求、审查制度以及历史流行病学等一系列文化基因组学(Culturomics)现象的发展<sup>[11]</sup>。也有研究采用博客数据进行分析,他们对某博客网站上“911事件”前后2个月的1084名用户的博客数据进行了语言学分析,发现在“911事件”发生后,用户的消极情绪增长、更加认知性和社会性地卷入、他们的语言中心理距离也变大,但2周之后,其情绪和社会参照便回复至基线,在接下来的6周中心理距离也回复至基线<sup>[12]</sup>。手机短信也是一种数据搜集方式,一项研究分析了50万条“911事件”当天的手机短信,发现愤怒情绪是最先开始发生的主导情绪;而焦虑虽然会产生,但会回复至基线,悲伤则是较后才产生的

情绪<sup>[13]</sup>。社会心理学家丹尼尔·吉尔伯特(Daniel T. Gilbert)及其学生<sup>[14]</sup>开发了一个IOS平台的手机App,让人实时报告其当下的思维、情绪以及事件。他们搜集了83个国家,数千名被试从18到88岁的人的86类活动。他们发现,最快乐的事情是性爱、锻炼、聊天、游戏、听音乐等,而相比于不幸福的事件,幸福事件通常是那些人们比较专心去做的事情,这与传统调查研究的数据结论<sup>[15]</sup>和福流(Flow)理论<sup>[16]</sup>的假设也较为相似。

我们看到,心理学家已经开始使用多种平台获取大数据信息,并采用语言学理论和数学方法进行了一些研究。但是我们认为大数据并不只是心理学数据采集方式的改变以及统计分析方法的数据驱动化,它并不等同于社会媒体,它也并不只是研究被试对象的增加。大数据同时给心理学理论和心理学研究的思维带来了变化。

## 1 方法革命对西方心理学体系构建的影响

托马斯·库恩(Thomas Samuel Kuhn)<sup>[17]</sup>认为科学的演进和发展以范式转移(paradigm shift)为轴。虽然范式还包括理论与观念,但通常是以方法的革新开始的。心理学致力于研究人类心理历程与行为,在古希腊时代已不乏对人类心理的探讨,但直到第一个心理学实验室的建立才被视为科学心理学的诞生。心理学实验室改变的是对人类心理的探讨方式,而并非是心理学研究的问题。当然,在微观层面上,新的研究方法会带来新的研究问题;但是在宏观上,心理学依然在探讨着人性与人类心理的本质。方法的革新并非给学科带来了研究问题上的本质变化,但是它却规定了一个学科的规范,带给了一个学科特定的研究渠道和根本的思维方式。大数据背景下的今天,心理学同样不会改变其心理历程和行为的主题,但是它会给心理学带来不一样的学科思维方式和研究手段。这种自下而上的影响必然影响到整个学科如何去思考问题、如何去探究主题、如何去执行研究。

心理学家艾宾浩斯(Hermann Ebbinghaus)曾说过:心理学有长久的过去,却只有短暂的历史。但是无论如何,西方在19世纪末建立起了科学心理学及其体系。实验方法的诞生使得心理学思想成为心理学,是方法的革新规定了这个学科的学科体系。大数据时代已然来临,哈勃空间望远镜一天的数据采集量甚至超过了天文学家之前两千多年采集的数据,

在这个意义上,说这是千年未有之大变革也并不为过<sup>[18]</sup>。现在,大数据背景下的心理学又重新面临一次方法的革新,这又是一次机遇和挑战,使我们中国心理学家有可能重新审视心理科学的学科体系。

心理学在我国出现很早,就像西方心理学家可以下笔皆溯之以古希腊一般,我国心理学家在心理学思想上甚至同样可以言必称孔孟老庄。从春秋战国时代至今,历代思想家均留下了他们对于知虑、志意、情欲、智能、才性等心理学观点的思考<sup>[19]</sup>。但很遗憾,现代中国心理学史在开始之初却是西学东渐的结果,是由留学西方的学者谱写的<sup>[20]</sup>。这样的好处是使我们快速地建立起了这样一门学科,但是坏处却是实际上并没有形成中国自己的心理学体系,甚至都没有心理学的中国化过程,现行的心理学体系和研究对象是完全西化的。文化心理学家彭麦克(Michael Bond)<sup>[21]</sup>曾感叹,如果心理学是由中国人创建的,那么它一定是一门强调社会心理学的基础学科。心理学并不是绝对而纯粹的自然科学,诸多心理学主题甚至是哲学问题,心理学科的社会性使得完全以祛魅(Disenchantment)<sup>[22]</sup>的现代科学甚至是生物学的研究方式来进行研究,丧失了心理学的独特性以及人心的超越。认识自己作为一个人的独特性不仅仅是根本的哲学问题,也是心理学的根本问题。由于对问题解释的多样性和不同角度,心理学可以如它的哲学母体般表现出不同的体系和形态。彭麦克的话说明中国的心理学可以是不同于西方心理学体系,用中国人的独特认识论来讨论心理学问题。当然,一些本土心理学家长期呼吁做有中国特色的心理学主题和内容,但是在当时心理学中的文化革命的背景下,他们的努力却止于进行一些本土性概念的探究,缺乏建立起一种独特中国心理学体系的魄力。当然,我们欣赏中国学者对心理学构念的不同解释,但是我们绝对不是以为了去寻求和西方的不同去定义中国心理学体系的重要性。相反,在大数据背景下,我们认为正是这样的时代背景与中国人独特的认识论方式给予了建构一个独特的中国心理学体系的可能。

自从20世纪末以来,心理学连续遭遇变革和范式转换。文化、进化、情绪和神经科学交替成为热门主题。在神经科学浪潮中,中国心理学家适时抢占了潮头,贡献良多。神经科学家对心理现象的解释体系并没有脱离西方,甚至是和西方无异的。当然神经科

学的震荡,有导致心理学覆灭的危险,一些心理学系甚至冠以脑与认知之名,一些认知和神经科学家从心理学系跳出而独立成群<sup>[23]</sup>。在这样变革的今天,大数据时代的来临,带给心理学特别是中国心理学一个整合、思考和重塑的机遇。

西方人强调个性,古希腊哲学家德谟克里特<sup>[24]</sup>以原子论的自然观来看待世界,认为世界是由每个不同且不可入的原子而构成。所谓的不可入,便是强调个体和个体之间的相互独立性,每一个原子都是一个个体。这种观点反应在对社会看法上,便是强调个体意识。而这种观点反映在其认识论上便是以还原思想为基础发展的现代科学研究。典型的自然科学物理学、化学、生物学均采用这样的分析手段。物理学研究各种粒子,试图找出最为微小的粒子结构;化学将物质还原为元素;而生物学则以细胞为基本单元作为研究起点。这种研究思想并未超脱朴素的德谟克里特的原子论。西方心理学同样基于还原论的体系,而且在神经科学革命时达到顶峰。冯特及其学生铁钦纳(Edward Bradford Titchener)开创了现代西方心理学的第一个流派,称之为结构主义(Structuralism)。铁钦纳意欲将心理的意识经验分为类似感觉、情感、意志这样的元素,用一种类似化学的方式来解析人类意识。虽然每个人都是由元素和细胞组成,但是我们不能仅仅只将每一个个体当作一堆元素或者细胞。格式塔(Gestalt)心理学对这样的还原思想提出了异议,认为整体不等于部分之和,而且我们的知觉原则是整体性的,我们能够将一条连续的散点知觉成一条线,也能够将照片上远近不同的个体知觉为同等大小,虽然他们在视网膜上的像是有大小的。但是比较遗憾的是,德国的格式塔心理学毕竟没有西方主流心理学中起到太大的冲击。威廉·詹姆斯(William James)的功能主义(Functionalism)心理学虽然不再强调意识的分解,转而认为意识的不可分性,强调意识的流动特征,但是詹姆斯的心理学依然是还原论的心理学。他强调自我,将自我分解为主我(I)和客我(me)。这种从“认识你自己”开始的思想本身就是个体意识的展现,强调这句话的古希腊哲学家苏格拉底宁愿赴死而拒绝他人的营救,这是因为苏格拉底将其自我作为了其行动的道德根据,将其自我意识极大化了。

之后的西方心理学流派皆沾染了还原论与个体主义的痕迹。尤其是认知心理学的计算机隐喻,彻底

将人作为一种高级生物低级化,表现出了一种很强的还原倾向。计算机是一种拟人化的产物,反而将人类心理还原为这样的拟人化物质表现为了一种逻辑的倒退,但是却方便进行实证研究。神经科学则将这还原推向了极致,将人还原为生物,还原为最基本的生物过程。当然,我们必须承认的是生物过程确实对心理现象有着直接的影响,且影响巨大。但是我们也不得不指出,人类心理确实不仅只是生物过程。是否任何心理过程都能够还原为这种生物或者物理的过程尚存争议<sup>[25]</sup>。用神经来解释诸如道德、自由意志这样的相对高级的人类心理现象似乎并不如用它来解释注意、记忆等相对低级的心理现象那样具有优势,因为其争议颇大<sup>[26]</sup>。

当然这可能是因为理论或者技术尚未发展到某一程度,对心理现象的生物生理机制的强调可以为心理学提供有力的研究方法和手段,但至少生物过程或者物理过程现在还不至于代替或者颠覆心理历程成为心理学的主要研究对象,虽然这是许多认知神经科学家的美好愿景。作为讲述心理学基本原理的教科书,其目录及顺序通常表达了这一学科的学科体系。西方心理学教科书一般遵循从神经元、基因等还原性的生理基础讲到感觉等最初级心理现象,再逐步推进到社会文化等主题的从低到高体系。重视前者,以前者为根的还原倾向明显。从低到高,强调低级过程对高级过程的影响,这是西方心理学的研究进路。中国心理学的教材也同样亦步亦趋。

但是西方人并非没有集体意识,只是集体意识是个体意识的外化,以个体意识为先,它变成了一个第二性的概念<sup>[27,28]</sup>。苏格拉底承认自己认同的法律,无论此法律判处是否公正,是否合理,他选择服从,这并不是因为他服从了法律这一集体契约,而是因为他强调了自己的个体意识。他并不是为了集体精神或者民族大义这样的东方传统理由而慷慨赴死,他的死是因为他的个体意识。但是西方文化的个体意识外化产生集体意识的方式容易产生异化。苏格拉底自己认同的法律导致了死亡,他所认可的规则反而成为了逼迫他自己的压力。个体意识走向异化是必然的结局。西方心理学同样面临这样的困境,以个体意识为先的还原论心理学相信以最基本的神经细胞、基因或者大脑活动便能够解释所有的心理现象。这样做的结局现在已经开始显现。强调基本生物过程的心理学因为强调基本的生物过程,将

自己逼到了绝境。越来越多的生物学家用生物学手段开始探究心理现象,相信生物学能够解决人类的心理问题甚至是社会心理问题。心理学科的疆域日益减少,它与生物学的疆界几近模糊,心理学系中的生物学倾向日益明显。传统社会心理学家、发展心理学家和临床心理学家与认知神经科学之间隔行如隔山的趋势越来越大。虽然有学者预测,到2038年,认知科学中占据九成江山的是心理学<sup>[29]</sup>,但是以此势头发展,不用20年,也许心理学已经名存实亡。这便是以个体还原论为基础的西方心理学产生的异化。强调最基本的生物过程却被生物学所钳制,以生物学为纲却沦为被生物学所代替,这是西方心理学异化的必然结果。而现今便是这一异化的高潮。异化代表着否定,在异化背后隐藏着的是另一次轮回,如果心理学仍然遵循西方传统,那么结果还将是再重复进行一次同样的发展,接着异化,接着再进行一次轮回。这也是为什么有学者认为心理学的理论和方法从来都在循环往复地发展的原因<sup>[30]</sup>。当然,异化不应该成为评估某一科学体系是否合理的关键标准,但是却深刻影响了学科的发展。西方心理学的异化发展至今实际已然大半去除了心理学中的心理成分,模糊了“心理”的概念,这已经开始钳制了学科发展。

## 2 大数据背景下中国心理学的体系

中国文化和西方不同,我们不是强调个体意识的文化,我们使用群体意识来定义个体意识<sup>[27,28]</sup>。儒家思想的伦理是建立在君臣父子的关系之上,“仁”实际上也是在阐述关系。这一点文化心理学中集体主义<sup>[31]</sup>以及相依自我<sup>[32]</sup>的研究均给出了准确的注脚。在实证上,如果让中国人更多地用“我是……”来描述自己,他们会更多地进行“我是中国人”、“我是我妈的女儿”、“我是某某的妻子”这样的具有群体中个体的表述方式<sup>[33]</sup>。而更为重要的是,在认识论上,中国人并非是还原论的,相反,我们是整体论的。围棋的全盘把握与中医的辨证论治都证明了中国人的整体观认识论<sup>[34]</sup>。我们认为世界是流变不居、充满矛盾但又普遍联系的<sup>[35]</sup>。我们能够分散注意的焦点而看到图画背景<sup>[36]</sup>,我们能注意到背景的变换<sup>[37]</sup>,我们能加工矛盾的信息<sup>[38]</sup>,我们倾向于场依存而不是场独立<sup>[39]</sup>。这种普遍联系的整体观恰恰符合大数据样本即总体以及从数据中发现未知联系的特征。如果说,神经科学变革将心理学推至极端的西化和还原

化,那么强调总体的大数据来临则给了按照整体观认识世界的中国学者和中国心理学以机遇。

我们将群体意识内化为个体意识,所以如果以中国文化为纲来构建一套中国心理学的体系,那么这一体系不会走向西方心理学那样异化的道路,成为反对和钳制自己的压力。我们只会在这—构架基础上不断积淀<sup>[40]</sup>。对比西方心理学体系,这是一个更为安全的心理学体系,而现在正是处在西方心理学异化的终点,和大数据整体思想的碰撞,中国心理学应以此为契机,建构一套独有的心理学体系。

那么以集体意识为先的中国心理学应该从何处作为起点和基础?维持集体的纽带在于纲常,在于共享的规范、道德和价值体系,这就是文化。个体的心理从本质来讲不光在反映我们的思维、情感和心理学的影响,它更是人类文化承载的载体。因此,我们认为,新的中国心理学体系的起点应该是文化。大而言之,这也是精神追求,文化赋予我们意义。研究已经证明,文化会影响个体的基本认知过程,也会影响人们的归因、分类、决策等社会认知过程<sup>[41]</sup>。在文化的框架下,道德也是中国文化强调的内容。孔子强调“礼”,认为人们首先得尊重礼法。而人的完满状态不过是“七十而从心所欲,不逾矩”(《论语·为政》),所以即使我们磨砺一生,最好的状态仍然是在规范中相对自由。而且这种在规范中相对自由的状态甚至成为了我们的精神追求。甚至还包括美、绝对意志、超越等主题,这是人心特别是中国人的心理历程作为一种可被研究的对象的起点。文化和精神追求作为一种形而上或者形式性的集体因素时刻影响着人类心理;而接下来的便应该是作为一种实体和经验性的集体因素的影响,我们将其称之为制度和社区的影响。而身处于集体中的人,其受到的社会情境影响在某一特定事件上要远远大过这个人的个体差异特征<sup>[42]</sup>。再者,无论是整个人类历史积淀的集体无意识<sup>[43]</sup>,还是个人历史包括其早期经验<sup>[44]</sup>都鲜活地将其烙印打在了个人之后的行为上。除去集体因素形而上与经验性的影响,剩下的第三部分,便是个体与集体的交互作用以及个体自身的因素上。自身因素这是西方心理学强调的因素,但只是新的中国心理学体系中较小的心理过程。这包括个体的特质等人格结构和需要、动机人格动力以及它们和情境的交互作用<sup>[45]</sup>。同时,也包括个体情绪、理性与行为问题,这是蕴藏于个体身上较为古老的话题。因此,我们认

为,中国心理学的学科体系可以遵循天、地、人相互关联的构建原则。从研究文化与精神的心理表征,到社会生活情境的制约影响,再到个体的知、行、情、理、德的身心体验。真正达到孔德(Auguste Comte)<sup>[46]</sup>所描述的人类最后科学的境界,即“具有社会特性的生物个体在文化情境中的行为”,这就是我们提倡的中国心理学。

既然我们重新架构了心理学体系,那么大数据背景下的中国心理学应该研究些什么?当然,如果我们将大数据仅仅看作一种方法论,那么我们可以和神经科学革命一样,将大脑活动作为一种观测指标,简单而地毯式地将所有先前心理学研究和教科书神经科学化,再在神经科学水平上或者大数据水平上进行理论整合。但是如我们前面所述,大数据带来的并非只是数据量的增加与计算方法的先进化,它实际上蕴含了整个学科思维方式的改变和学科体系的倒转。我们认为,按照新的心理学学科体系,大数据背景下的中国心理学体系可以研究以下内容:

(i) “天”的意义。人类社会创造的文化和精神对人们心理的影响。比如意识问题,如群体中的人理解的意识是什么、中国人的自由意志观等问题;文化问题,如使用大数据方法进行文化比较、了解文化的进化、破译文化基因的遗传密码等问题;信仰、善与美的问题,如信仰对人类行为的影响、道德思维、具身道德、美德、良知、跨文化审美认知等问题<sup>[47,48]</sup>。

(ii) “地”的制约。人所创造的社会制度与社会关系对人性的制约及规范作用。比如在大数据背景下进行一系列诸如文明历史及意义发展、个人历史信息预测等历史问题;家庭、学校、社区等社会结构问题;法律、政治等社会制度问题;亲属、同事、朋友、伴侣等社会关系问题;社会情境分类问题等<sup>[49,11]</sup>。

(iii) “人”的活动。人类心理活动和过程的身心体验。比如在大数据背景下阐述人格结构与类型、智力和创造力的本质等个体差异问题;情绪和理性对行为的影响、动机和需要对行为的驱动等行为动力问题;知识的本质与储存方式、知行合一的心理学表达、学习的实质等教育和发展问题;与人心相关的身体体验、生理过程与生物学结构等问题。

### 3 中国心理学新学科体系下的大数据心理学特征

就像我们所讨论的那样,实验室实验方法使得

现代心理学建构起了其西方体系，而大数据的方法变革使得我们有可能建立起符合中国文化和中国人认识论的心理学体系。当威尔海默·冯特(Wilhelm Wundt)建立其心理学实验室时，他和整个学科面临了诸多的问题。不得已与退而求其次的委曲求全方法实际上在心理学思考中大量存在，由于长期的使用，这些因为不得已而使用的方法变成了经验标准甚至学科规范。做一个简单的反事实思维，如果冯特能够很容易、很快地测量所有人的感觉，我们有理由相信他会去这样做，而不是仅仅依靠几个研究被试的结果。当然这样说缺乏实证精神，历史无法操纵，但是我们相信这样的反事实思维是合理性的。心理学思维和方法上的不得已至少表现在：第一，由于无法搜集或者无法很容易、快速地搜集到全体数据，改用样本来通过统计方式代替全体；第二，由于无法经济地获取个体历史数据，建立在概率统计基础上的心理学结论能够发现普遍规律但无法将其用之于个体；第三，由于无法很好在现实生活中观察想观察的现象，心理学实验方法必须在实验室中创设各种非真实情境；第四，由于心理过程和行为的量化缺乏统一标准，心理学构念必须由研究者根据其个体经验将之转化为可以操作化的概念并代之以测量。这些不得已而为之的方法已然在心理学研究者的思维中根深蒂固，它影响着心理学的实验设计、结果推倒与研究评价。大数据背景下，许多问题已经不需要退而求其次。心理学需要改变的首先是破除许多所谓的学科规范和经验标准，在变革的时候不需要固执于本来便是委曲求全的方式，换之以本来就应该如是的方式来对待心理学问题。这些大数据心理学特性与思维方式表现在：

### 3.1 样本即总体

由于不经济的原因，心理学家在以前不大可能获得总体数据。人格和个体差异研究者经常使用心理测量方法，因为量表项目和统计方法的关系，他们还有可能去选取超过一千人的大样本<sup>[50]</sup>。而社会心理学家因为使用实验室实验方法，所以分派在每个条件下的研究被试通常都为几十人。发展心理学家或者临床心理学家甚至可能使用更少的样本。但是除了对单个被试进行单纯干预的研究，心理学的主旨在于揭示普遍规律。所谓普遍规律意指将人类作为一群体而适用的规律。很明显，要一个群体中非

常少的一些个体的规律来替代整个群体的规律，是需要有先决条件。其一是，这个群体中的所有个体没有任何区别；其二是，通过科学地选取方式，我们有可能使得选取的个体能够代表整个群体。

认知科学家特别是认知神经科学家通常会暗含第一个假设。他们将人脑看作计算机的处理方式通常便是如此假设的。我们对外界信息进行编码、存储、加工和提取的过程似乎是普适的。因此研究基础生物过程和认知过程的心理学家通常使用很小的样本。但是我们可以轻易地看出这种假设和基于这种假设的研究结果之间的不协调性。首先，即使在最基本的认知过程上，仍然会存在个体差异。不同的人会注意不同的信息，也会记住不同的信息。每个人甚至在工作记忆的容量上也存在区别。其次，根据第一个假设，群体中的个体并无差异，那么我们完全有理由相信不同的人在面对同样的刺激时会做出一模一样的反应。当然，这个假设更应该解释为如果在同样的时空条件下，如果每个被试有同样的历史经验，他们会做出完全一样的反应，就像同样配置的计算机计算过程和呈现的结果无异一样。但是我们会在诸多fMRI研究中，发现分派在同一实验条件下的研究被试做出了不同的结果。也就是说，在大部分fMRI研究中，每个实验条件下结果的离中趋势并没有理论上的那么小。这本身就是一个逻辑错误，部分心理学家根据他们信奉的前提条件进行研究，而研究的结果恰恰否定了这一前提条件。

如果我们承认个体间确有差异，那么小样本能代表总体的可能性只有随机抽样这一条路径可以选择。确实，如果遵从严格的随机抽样程序，无论是何种程序，我们在数学上是可以推测总体的全貌的。类似盖洛普之类的调查实际上已经部分证实了其效度。但是，大量主流心理学研究却并没有遵从随机抽样的程序。这并非心理学家不想随机抽样，而是在之前的历史背景下通常做不到。心理学家能够做到的只能是随机将并不随机的样本分配到不同的条件中去。随机分派在统计上能够解决一些无关变量的影响。但是我们依然无法因为几十人或者上百人的实验得出人类的普遍规律。而且，基于随机抽样的原则，要抽取一个普遍适用于人类整体的样本，是不可只有几十或者上百人的。为了能够适应统计规律，有时候心理学将30人便称为大样本，这实际上是囿于样本的限制而不得已去适应研究与分析方法，而没

有考虑研究假设和结论的无奈之举. 基于样本的研究其结论严格来说只能用于这一样本, 再向外推广, 其概括性和生态效度是值得怀疑的.

有经验的研究者都会知道绝大多数心理学研究是基于西方的(western)、受教育的(educated)、工业的(industrialized)、富裕的(rich)和民主的(democratic)社会及其人群的. 大部分主流心理学研究结果是基于美国大学生样本, 研究者根据其方便来取样的<sup>[51]</sup>. 有学者甚至严肃指出心理学权威研究杂志《人格和社会心理学杂志》(*Journal of Personality and Social Psychology*)应该更名为《美国大学本科心理系学生的人格和社会心理学杂志》. 这在人群中仅占5%的样本毫无代表性可言<sup>[52]</sup>. 他们并不是所谓的“普通人”, 他们在视知觉、公平、合作、空间推理、分类和演绎推理、道德推理、自我概念及其相关动机甚至是智力的遗传性上都和其他95%的人存在区别<sup>[53]</sup>. 因此, 将某一并不随机抽取的人的结果作为某一特定行为或者心理现象的基础, 这种先验标准是错误的, 也是不合实际的. 大数据背景下的心理学实际上模糊了样本和总体的概念. 样本大到就是总体. 也就是说现在我们可以有条件获取所有人的数据, 而这样的结果实际上可以推论到所有人.

为了掩盖样本造成的问题, 传统心理学家在方法和理论上都给出了其辩解. 在方法上, 心理学家采取考察尽可能多的样本的方式来验证某一特定结论的信度. 一种方式是进行元分析(meta-analysis), 即用手工的方式选取尽可能全面的研究结果, 用数学的方式来综述, 看对同一问题的不同被试与不同方式之间效应如何. 元分析方法极度依赖已发表的文献. 但是已发表文献同样存在巨大的问题. 已发表文献通常是结果显著的研究, 而事实情况则可能是这一结果并不真实存在, 但是由于某些特定原因, 某些特定研究者得出了显著的错误结果, 而他的研究得到了发表, 导致的后果是让学界认为这一结果是真实的. 也许十次实验只有一次显著的结果, 而这一次得到了发表, 那么学者们实际上并不知道这个实验的结果应该是失败的. 这种在心理学界普遍存在的发表偏差(publication bias)<sup>[54]</sup>实际上极大影响了元分析的有效性, 也极大地影响了学者们认识真实心理规律. 第二种心理学家常用的多样本交叉验证研究信度的方法就是重复研究. 可重复是实验科学的基本特征, 是验证研究结论信度的不二方式. 但是由于

发表偏差以及对原创性的追求, 验证研究在心理学界并不盛行. 在诺贝尔奖得主丹尼尔·卡尼曼(Daniel Kahneman)<sup>[55]</sup>公开批评心理学启动研究范式, 并号召启动研究进行重复验证之后, 加之因实验数据丑闻频发而导致整个心理学界形象告急的时候, 一些SCI或者SSCI索引杂志抛开原创性要求, 开始声明接收重复验证性文章. 美国心理学会(American Psychological Science, APA)也专门出版杂志接受重复性验证文章. 但在大数据背景下, 这种影响是可以得到轻易解决的, 一者是基于大数据的心理学研究由于其样本便是整体, 并不需要过多的元分析; 二者是, 即使是基于传统的研究方式, 在理想状态下, 大数据背景完全可以将心理学的各种研究数据都做存留, 动态性地报告出某一特定心理学假设的实时元分析结果.

在理论上, 这些基于样本的研究在某些情况下还有辩解的余地. 心理学的基本原则便是, 心理历程和行为受多种变量影响. 一条心理规律必然有着诸多的限定语. 在实验中, 除了主效应的发现, 实际上, 心理学家更执着于找到一个调节变量(moderator)或者说两个变量之间的交互作用. 而被试的限制实际上和心理规律的限定于有重合之处. 如果美国心理学家仅仅关系美国中产阶级白人的心理规律, 那么他们选择WEIRD的研究被试并没有错误之处. 但是, 实际上只要我们承认人与人之间是存在个体差异, 那么限制就可以在逻辑层面无限分割于每一个个体, 导致我们无法得到普遍规律, 只能得到个体规律. 这显然并非心理学家所愿. 大数据的出现, 不仅是将样本变成了整体这么简单, 而且将这一悖论得到解决. 由于全体的数据量, 我们可以轻易地在任何层面上获取数据、得到结论, 可以方便地得到有任意限定语的结论. 在得到这些普遍规律的同时, 我们也可以不放弃个体的个性规律.

### 3.2 个性即规律

心理学家试图得到适用于人类全体或者某一群体的普遍规律. 我们已经讨论过, 适用于全体的普遍规律在传统心理学中很难得到. 即使得到普遍规律, 另一个问题是, 它无法应用于个体. 亦即, 我们得到了普遍规律充实了学科知识, 但真正能够指导自己生活的心理学结论却在理论上是没的. 当然, 群体规律能够运用于群体, 对于政策制定者或者思考社会的智者来说, 这样的规律能够让他们迅速把握总



体样貌。但是,对于绝大多数生活在自己独立生活中的个体来说,他们关心的还是自己作为一个个体的生活。那么,当心理学结论运用于单个个体来说时,便会产生问题。

社会心理学的研究发现,带上墨镜能使人行行为更不道德<sup>[56]</sup>。这是对某个群体的研究结果,即使这个群体是具有高度的代表性的,也不能说明任意一个人带上墨镜,他就会变得更加不道德。这是因为心理学的结论是建立在一个群体的平均水平上的,我们只能说:平均来说,带上墨镜会使人更多地进欺骗。研究表明,写宽恕信能够有效提升个体的幸福感<sup>[57]</sup>。但是,这一结果也是基于实验比较的。并非所有写了宽恕信的被试其幸福感的提升都高于没有写宽恕信的被试,而只是写下宽恕信的人其幸福感提升的平均水平要高于控制组的人。但是,实际上,无论是心理学家还是媒体宣传都容易犯这样的错误。我们能够根据心理学的研究结论给出普通个体以建议,但是,建议只是建议,群体结论是不可能适用于每一个个体的,严格来说,这样的建议犯了逻辑错误。

探寻普遍规律的传统心理学无法避免这样的逻辑错误,因为它的研究结论并不是针对个体或者某个特定的人的。普遍规律是学科知识积累和探索的需要,而对个体生活以指导则只是附带产物。即使以指导个体生活为己任的传统心理学家,其研究和指导实践之间也必然存在着这样的逻辑错误。但是大数据心理学家却可以在探寻普遍规律的同时,比较方便地给予个体以指导。以亚马逊(Amazon)公司为例,由于它拥有用户的历史搜索与查看数据,其推荐列表能比较准确地根据用户偏好给出推荐商品<sup>[58]</sup>。这种推荐的方式是针对个体的,因为每个个体都有不同的历史记录。而其推荐何种商品却不是自上而下的理论式的,仅仅是一种数据驱动的计算方式,根据别人的搜索历史来模拟你的偏好。这种自然式方法反而更能够运用在具体的个体生活中。

传统心理学家并非不想对个体进行预测,但是获取每个个体的历史数据近乎不可能。在大数据背景下,获取这样的数据已经可以比较轻松。根据历史数据,我们能在大数据背景下对每个人个体进行独特性的分析,给出个体特有的规律,从而使个体运用心理学结论成为可能。假使传统心理学家能够得到个体数据,那么他们所进行的预测,也多半会基于理论,或者自上而下的加工方式。比如人格心理学家会

根据个体的个体差异特征来推断其之后的行为。但人格特质本身便是一个集合性的平均概念,它无法进行预测,它只能在平均水平上进行探讨。一个外向的人并非不会外向,他只是在所有行为的集合上表现得外向而已。由于心理学的这种解释方式始终还是建立在以平均和普适性为基础的心理学术念上,即使存在个体历史数据,其对个体行为的预测也始终不能令人满意。而大数据背景却使得心理学家能够动态化地根据所要研究的心理现象和行为在自然无序状态下的结构,根据大量他人在类似历史条件下的真实行为来推测某个特定个体的心理现象和行为。大数据的革新并非只是数据量的增多,实际上还蕴藏了理论建构方式的变革。

我们有理由假设,在心理测量学中,由于大数据的出现,其基本理论和验证方式能够发生改变。简单地说,由于大数据背景下个体历史数据的搜集成为可能,那么从传统着重整体测验及其项目间关系的经典测量理论(Classic Testing Theory, CTT)到着重个体项目特性的项目反应理论(Item Response Theory, IRT),心理测量理论的着重点是否会发生变化?计算机自适应测验(Computerized Adaptive Testing, CAT)的产生是为了将个体作为测验的参与主体的特性突出,而不是将测验的重点集中在测题上。但是由于理论的缺乏,以个体为中心的测验始及其理论体系始终没有建立。人格和个体差异的复杂性告诉我们,不同的人在面对同一套测验体系或者同一个项目时进行的理解不会是完全同质的。富裕的个体和贫穷的个体在理解“金钱”这个概念时必然不同,他们对同一个涉及此概念的测验或者单个项目的反应就不能在同一个算术体系下进行平均和比较。因此,建立一种以个体为中心而非以测验为中心的心理测量理论是必要的,而大数据的背景则为这种理论的形成提供了条件。同时,个体历史的保留同样可能解决某些社会问题,如对社会事件中涉及的个体的全面把握,以及类似高考这种不顾历史数据而定于某次抽样的选拔方式的讨论。

一个测量的好坏在于其信度的检验方式。即使使用传统测量理论,大数据背景也将进行革新。测量是数据驱动的,大数据背景会改变心理学对于理论特别是以数据驱动的理论的检验。由于数据的更新保存,对某个测验来说,常模可以说动态的,因为它不断在进行补充和更新。那么一个好的心理学建构,



在这个意义上来说,它一定是一个基于数据而动态变化却又始终保持结构稳定的理论建构.如果它是一个房子,那么房梁地基都不会发生变化,而随着数据的变化住在其中的人却不断变化,但无论人怎么变,屋子总是那个屋子.以大五人格理论为例,如果他是一个好的理论,那么它在数据增加的时候,或者随着时代变迁,那它的五个因素就不会发生变化,但是五个因素对应的项目却可以发生时代的变迁.这同时也表明,一个好的心理学建构它始终是以数据和计算为基础的,却又是能应用于数据和计算的.理论和数据的关系是双向的.是数据和计算组成了理论建构,而这一建构同样能够预测一群数据中某一个点的属性或者行为.就传统心理测量观来说,如果说大数据或者动态常模有颠覆,那么它实际上可以不颠覆经典测量理论甚至是项目反应理论的观点,但是它必然颠覆了经典测量理论和项目反应理论的验证方式.对于经典测量理论来说,一个构念的结构应该在一个动态的常模上有着动态的稳定性;而对于项目反应理论来说,一个项目的优劣也在于它应该在一个基于动态的被试选择中保持其项目特征曲线基本参数的稳定性.

就大型的心理学理论来说,大数据背景下动态的数据驱动理论或许要强于至上而下式的是思辨理论.因为思辨理论常会指责数据驱动理论是描述性理论,而它们可以做规范性的事情.即数据驱动的理论容易犯下从“是”推论出“应该”的自然主义谬误.但实际上,我们可以说,一组动态数据驱动的理论已经将“是”极大化了,把这种“是”已然推向了“应该”也就是能够预测行为的边缘,即它虽然只是描述了自然状况下各种心理和行为的关系,但是像我们讨论过的那样,它能够根据这些自然的数据关系,做出较为良好的预测,正如前述亚马逊的例子一样.

### 3.3 情境即实验

心理学研究最重要的方法便是实验方法.心理学实验室的建立是科学心理学诞生的标志.在我们看来,区别于其他方法,实验方法的特点便是不必真实.这是优点、也是缺点.心理学实验方法不真实的优势表现在:第一,不真实意味着非自然条件,即可以进行任意操纵,这极大增加心理学的研究速度.研究者在非自然实验室条件下可以创设任意实验条件,让被试身置其中.这样便可以考察任意研究者想考

察之现象,而不必像自然观察法那般必须等待自然条件中想考察的现象自然出现.这样能够迅速积累实验,是心理学科得到快速发展,也增加研究者的控制感与对研究的把握.第二,不真实意味着在非自然条件下可以极大剔除自然条件下的各种混淆变量,尽量只考虑心理学研究中自变量与因变量的关系.心理现象和行为受诸多因素混合影响,而实验室情境可以通过精确的控制,使得实验组和控制组在除了被操纵的自变量这一个条件不同外,其余条件均保持相同.这样,实验组和控制组之间的比较才反应了自变量对因变量的影响.而这种除了自变量之外的一致性,在自然条件下几乎不可能存在.虽然可以通过如协变量等统计方式来矫正混淆变量的影响,但是其精确程度以及因果推论的准确性与实验室条件不可同日而语.因此,心理学实验室实验的不必真实特点使得这门学科在短期内能够积累大量知识,也使得心理学研究的精确化和可重复性大增.重要的是,正是实验方法使得心理学可以宣称自己是科学.不必真实的实验方法使得心理学家能够自豪地在教科书开篇有底气地论证心理学作为一门科学的合理性.他们可以放心地说,一门学科的科学与否在于研究方法,而不是研究主题.某种程度上来说,正是实验方法的科学性赋予了心理学以科学的特征.物理、化学教科书在开篇之处不会花大力气说服读者认为他们是科学,但是几乎所有心理学教科书都在做这件事情.作为心理学家,我们必须承认这是一个学科的无奈和不自信.因此,传统实验室实验方法在心理学中的地位无可撼动.

但实验室实验方法是否就无懈可击?答案必然是否定的.其缺点仍然是由其不必真实的特性而引起的.就像我们前面所探讨的那样,心理学试图得出普适性的规律,而这样不真实的实验情境中得出的结论能否适用于其他情境是值得怀疑的.严格来说,从什么样的实验情境、什么样的实验操纵、什么样的实验观测方式、什么样的研究对象中得出的结论只能适用于同样的实验情境、实验操纵、实验观测方式以及研究对象.在研究的生态效度与研究的概括性上,实验室实验比较欠缺.

而大数据背景下的心理学研究,这一基本方式有可能得到扭转.我们并非要抛弃操纵自变量、控制混淆变量、观察因变量、比较实验组和控制组的研究逻辑.只是大数据背景下的心理学数据情境众多,而

且几乎是自然条件,不存在代表性问题。大数据使我们能够比较容易地选取不同的情境而进行比较。因此,在大数据中挑选情境便是实验方式。传统的自然观察法的弱点在于无法等待想要研究的现象出现。而在大数据中,实际上现象均已出现,心理学家需要做的只是挑选,而并非等待。这种自然,而且能够比较的新型实验方式,不存在研究结果概况性的问题,也并不缺乏实验精神和因果探究的逻辑。对比实验室实验的第一个优势,大数据背景下的心理学研究没有了操纵,但是在逻辑上并非不能比较,因为操纵还是为了比较。实际上为了进行研究,大数据的数据搜集者并非不能进行操纵和干预,抛开需要讨论的伦理问题,这是可行的。同时,对已有数据的挖掘完全能够使各种已有情境变成实验的自变量,从而进行比较。我们完全可以将大数据的这种方式看作自然观察的未来形态,即观察者已知所有他要观察和不需要观察的现象。这样的话,心理学研究便变得简便、经济、普适而且可比较。

对比实验方法的第二个优势,大数据方法确实存在无关变量无法控制的问题。但是大数据研究者一般忽略此问题,他们甚至宣称自己不注重细微的精确性。从统计思想上来说,随机抽样能够比较有效地控制干扰变量的影响。但是随机过程能够带来的最好效果只是模拟总体,在任何变量上的维度中随机样本和总体保持较为一致的稳定性。如果总体在某一个维度上表现出系统误差,那么样本同样也会保持系统误差。这是大数据无法避免的问题。但是如果从大数据的全体中抽取不同的情境进行实验性比较,那么通过比较也可以以类似实验室实验的匹配方式来解决这一问题(如A/B testing方式, <http://visualwebsiteoptimizer.com/case-studies.php>)。

实验室实验作为心理学主导方法的另一个原因是,实验是公认进行因果推理的最有效方式。心理学实验通过比较和控制的方法得出因果关系,这遵循了密尔(John Stuart Mill)<sup>[59]</sup>5种因果推论方法中的求异法。一般来说,心理学讨论两种关系,一种是因果关系,一种是相关关系。一些大数据研究者认为大数据需要探讨的就是相关关系,而不需要因果关系。诚然,从原理上来说,因果关系比相关关系多出了一些信息,即谁因谁果以及是否存在第三变量的问题。相关关系已然给两个变量之间的联系做出了解释,虽然这一联系并不一定是必然的。但是按照休谟(David

Hume)<sup>[60]</sup>的对因果关系的理解,如果我们知道两个变量之间的必然联系,那么剩下需要确定的只是空间是接近性和时间上的延续性。而后两者在大数据中是非常容易获取的。因此,大数据并非不能获得因果关系。再者,根据心理学的统计方法,进行实验统计的方差分析方法,在本质上仍然是考察操纵的变异与误差变异的检验,二者之和作为总变异。在数学原理上,这与考察相关关系的回归模式并无二致。因此实际上,真正影响了因果关系的还是实验假设,同样一套操纵的数据既能说明暴力电视节目增加攻击行为,也能说明更加攻击的人喜欢看暴力电视节目,真正影响因果关系的还是研究者的假设<sup>[61]</sup>。而且若进行暴力媒体与攻击行为的影响研究,在统计思想上,我们实际上只比较了观看暴力电视节目者的攻击行为和未观看暴力电视节目者的行为。而要做因果推论,我们还必须掌握观看了暴力电视节目者若未观看时的行为以及未观看暴力电视节目者若观看时的行为。而这两者,实际上再精确的实验室实验也无法了解。因此,实验室实验的因果关系在逻辑上也是不严格的。大数据方法把真实情境作为实验,在得出因果关系上,我们认为能够和实验室实验有几乎等同的效力。在已有研究中,大数据的主要功用实际上并非只是简单阐明了相关关系。大部分研究者运用大数据去进行预测或者填补并不知道的数据缺失。这两点功用实际上暗含了因果关系的功能在内。当然,纯粹以计算数学模型建立的理论在建立之时并没有考虑变量之间的因果关系,只是用数据驱动的方式建立起一个联系。但是,对这种模型的解释或者说对这种模型的使用,却并不只是相关那么简单。

### 3.4 数据即行为

在心理学诞生之初,意识的结构和功能才是心理学的研究对象。但是自从华生的批判,心理学才转而将目标关注行为。这一转变使得研究行为在心理学的研究对象中占据了最重要的地位。虽然一般心理学都会把心理过程和行为对应,但是一般的认知心理学都会用反应时等行为指标来推测心理过程。当然在神经科学研究方法中,神经等生物过程代替行为作为心理现象的反映指标。但是传统心理学仍然采用大量的行为实验来探究心理现象。我们的观点是,大数据背景下,数据就是行为。这可以理解

为,大数据记录的信息实际上就是普遍化的人类行为,而且这种数据行为具有自愿以及可普遍化两个特征。

首先,大数据作为行为由于大部分是自然条件发生的,它是人自觉自愿的行为。比如微博留言和评论,大部分时候虽然可能有自我展示的成分和动机在内,但是这样的数据至少是自觉的,也不是被强迫的。有经验的研究者都会发现,心理学的大多数被试是课程被试,他们进行实验是有获取学分或者达到课程要求这一目的在内的。而现今大量使用的类似亚马逊Mturk这样的网络数据,被试也是因为为了获取金钱而进行心理学实验的。这样的被试在得出结论的代表性上至少是存在争议的。而其动机通常也会在心理学实验中引出不必要的额外变量。即使不是为了钱或者学分的普通被试,那么这样的被试通常也以女性居多,他们在某些个体差异特征,例如利他程度上是和别的被试有区别的。大数据中的行为数据却基本是在自然条件下自然发生的,并不是受到实验者强迫,并将他们置于一个受控的环境中完成的,譬如越来越多的可穿戴设备以及量化自我技术(quantified self, <http://quantifiedself.com/>)使得我们能够很轻易地在自然条件中收集生理和行为数据。这是大数据行为数据的一大优势。

其次,大数据的行为数据是可普遍化以及可以比较的。在心理学研究中,由于无法直接观察心理现象而需要用行为去推测,导致可供观察的行为五花八门。在心理学研究中,对同一心理现象的观察可以采用不同的方法。每一位研究者都以创造不同的范式和方法为乐,而且新的独创的观察方法通常还会受到高档次杂志的青睐。这是学科的创新和理性之美,但是这更使得学科的系统庞杂混乱,许多相似概念之间没有可比性。例如研究攻击行为,研究者可以观察人们是否会对一个玩偶进行击打,也可以观察人们是否会对一个犯错误的人进行点击,甚至还可以看一个人会把他人照片撕得有多碎。不同的方法之间很难直接进行比较,而且每种观察或者用心理学术语来说,每一个操作性定义都不能完好地代表同一个概念。但是大数据较好地解决了这一问题。由于数据已经搜集好,研究者所做的更多事情是从已有数据中挖掘出行为,这个时候每个研究者在大数据中定义同一种行为基本会有差不多的操作定义,使得大数据产生的结果之间能够比较。

## 4 总结与展望

对于中国心理学来说,大数据是一个机遇和挑战。机遇在于大数据革命所蕴含之理念切合中国文化,中国心理学家理应首先受益,也因之可以更符合中国朴素思维方式化地对心理学系统重新进行思考。需要强调的是,这并不是说西方心理学就不会或者不适宜使用大数据的方式来进行研究,也不是说中国心理学就只能研究群体而忽视个体。心理学研究的历史已然表明并非如此。我们意在指出大数据革命可以使得中国心理学重新思考研究心理学的进路,以期改变西方个体化进路使得还原论极化以致心理学异化并生物学化而难以自拔的现状。符合中国朴素认识论的大数据为心理学学科进路的转向提供了可能,为心理学的发展与研究方向提供了新的选择。这一变革不会是中国心理学所独有,但论此变革背后之思想渊源却最为贴合中国朴素认识论。

而大数据革命又切实给我们带来了挑战。不难看出:首先,大数据改变的不是技术,而是心理与行为。诚然,大数据使得计算机信息技术有了更大的用武之地,但是技术的革新或者升级演进并不是大数据对人类所产生的最深层次影响。大数据最终所影响的是行为,是活生生的人,而不是冷冰冰的数字。人类的行为、社会的行为、政府的行为、国家的行为,大数据表面影响着浅层的信息技术,而深层影响着的是人类的活动方式、社会的变迁方向、政府的组织架构和国家的战略决策。这些全面的影响使得心理学家担负了更加重要的职责。其次,大数据不是方法革命,而是思维革命。大数据革命并不是要改变研究方法,而是要改变实施和设计研究的指导思想。在单纯技术领域,大数据也许会使得计算方法更加精益求精,但是在面对具体的问题时,大数据往往使得我们从另一个从来没有的角度来看待问题。比如样本即总体、个性即规律、情境即实验、数据即行为的特点,使我们能够站在更大的视野上来看待人类社会和行为,彻底改变我们看待问题的方式和设计研究的角度。快速适应思维转换乃是大数据变革下心理学家需要投入努力的地方。再次,大数据研究不是算法导向,而是问题导向。应该说,任何研究都不是算法导向,而是问题导向的。只是有时候研究者会迫于方法,而且解决问题的方式。在信息过剩的今天,能够从大量信息中提出问题,是如何利用精致的方法解决问题的前端环节。只有算法,而不解决实际问题

的研究是空洞的。发现问题、提出问题、解决需求的价值在研究层面上远远大于精致计算、细致操作的价值。在这样的背景下,心理学家不应纠结于具体的计算,而应在适应大数据思维的条件下重新审视心理学研究的问题,积极寻求与技术专家的合作。最后,大数据并不是学科改变,而是学科颠覆。如果一次革命并未对一门学科产生质的改变,那么这次革命是不成功的。大数据对于技术领域的改变不可谓不大,但是相对于心理学,前者的改变可能还只是量上。我们已经提出,大数据势必改变整个心理学或者保守来说,是中国心理学的学科架构。因为它迎合了中国人的内在思维习惯和外化行为模式。我们需要适应这样的颠覆,并将大数据的思维模式有机地融入朴素的认识论。可以想象,在不久的将来,也许连“大

数据”一词都将不再被提起,就像我们只有在停电时才会意识到我们身处一个电的世界,但我们不会时常提起它。那么当那一天来临的时候,大数据已经深入我们的生活。或者深入地说,是大数据的思维方式潜移默化地深入了我们的生活,是我们慢慢习惯了大数据的行为模式。大数据这一词的远去,留下的是一系列在未来看来可能并不算先进的算法和在任何时候都会熠熠生辉的思想。

大数据变革不仅让心理学家找到了一种新的研究方法和数据来源,更重要的是能让我们重新思考心理学的研究对象、目的及理论体系。创建中国心理学的理论,不是要终结我们对世界心理学的结合和学习,而是借助新的研究范式和新的发现,为世界心理学做出我们中国心理学家应有的贡献。

## 参考文献

- 1 Ginsberg J, Mohebbi M H, Patel R S, et al. Detecting influenza epidemics using search engine query data. *Nature*, 2008, 457: 1012–1015
- 2 Butler D. When Google got flu wrong. *Nature*, 2013, 494: 155–156
- 3 Cook S, Conrad C, Fowlkes A L, et al. Assessing Google flu trends performance in the United States during the 2009 Influenza Virus A (H1N1) pandemic. *PLoS One*, 2011, 6: e23610
- 4 Lazer D, Pentland A S, Adamic L, et al. Life in the network: the coming age of computational social science. *Science*, 2009, 323: 721–723
- 5 Watts D J. A twenty-first century science. *Nature*, 2007, 445: 489
- 6 Bollen J, Pepe A, Mao H. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In: *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media*, Barcelona, 2008. 450–453
- 7 Golder S A, Macy M W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 2011, 333: 1878–1881
- 8 Kahneman D, Krueger A B, Schkade D, et al. A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 2004, 306: 1776–1780
- 9 Seligman M E P. *Flourish: A Visionary New Understanding of Happiness and Well-being*. New York: Simon and Schuster, 2012.
- 10 Schwartz H A, Eichstaedt J C, Kern M L, et al. Characterizing geographic variation in well-being using Tweets. In: *Proceedings of 7th International AAAI Conference on Weblogs and Social Media (ICWSM)*, Boston, 2013
- 11 Jean-Baptiste M, Shen Y K, Aiden A P, et al. Quantitative analysis of culture: Using millions of digitized books. *Science*, 2011, 331: 176–182
- 12 Cohn M A, Mehl M R, Pennebaker J W. Linguistic markers of psychological change surrounding September 11, 2001. *Psychol Sci*, 2004, 15: 687–693
- 13 Back M D, Küfner A C, Egloff B. The emotional timeline of September 11, 2001. *Psychol Sci*, 2010, 21: 1417–1419
- 14 Killingsworth M A, Gilbert D T. A wandering mind is an unhappy mind. *Science*, 2010, 330: 932
- 15 Kahneman D, Krueger A B, Schkade D, et al. A survey method for characterizing daily life experience: The day reconstruction method. *Science*, 2004, 306: 1776–1780
- 16 Csikszentmihalyi M. *Flow*. New York: Harper Collins, 1991
- 17 Kuhn T. *The Structure of Scientific Revolution*. Chicago: University of Chicago Press, 1970 [库恩. 科学革命的结构. 北京: 北京大学出版社, 2003]
- 18 Mayer-Schönberger V. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013
- 19 Gao J F, Yan G C. *History of Chinese Psychology (in Chinese)*. Beijing: People's Education Press, 1985 [高觉敷, 燕国材. 中国心理学史. 北京: 人民教育出版社, 1985]

- 20 Yang X H. History of Psychology (in Chinese). Ji'nan: Shandong Education Press, 2000 [杨鑫辉. 心理学通史. 济南: 山东教育出版社, 2000]
- 21 Peng K P. Preface (in Chinese) in: Plous S. Decision Making and Judgment. New York: McGraw Hill, 1993 [彭凯平. 序. 决策与判断 (普劳著, 施俊琪 译). 北京: 人民邮电出版社, 2004]
- 22 Weber M. The Protestant Ethic and The Spirit of Capitalism, Roxbury: Roxbury Publisher, 1930 [韦伯. 新教伦理与资本主义精神. 上海: 生活·读书·新知三联书店, 1987]
- 23 Ivry R. Big data has left the station. APS Obs, 2013, 26: No 1
- 24 Deng X M, Zhao L. History of Western Philosophy (in Chinese). Beijing: Higher Education Press, 2005 [邓晓芒, 赵林. 西方哲学史. 北京: 高等教育出版社, 2005]
- 25 Baumeister R F. Do the Laws of Physics Permit Any Exceptions? Psychol Today Blog, 2009
- 26 Smith K. Neuroscience vs philosophy: Taking aim at free will. Nature, 2011, 477: 23–25
- 27 Deng X M. An analysis of psychological models of Chinese and western culture (in Chinese). J Northwest Norm Univ (Social Sciences), 2010, 2: 1–10 [邓晓芒. 中西文化心理模式分析. 西北师大学报(社会科学版). 2010, 2: 1–10]
- 28 Deng X M. An analysis of psychological models of Chinese and Western culture (in Chinese). J Northwest Norm Univ (Social Sciences), 2010, 3: 1–12 [邓晓芒. 中西文化心理模式分析(续). 西北师大学报(社会科学版). 2010, 3: 1–12]
- 29 Gentner D. Psychology in cognitive science: 1978–2038. Top Cogn Sci, 2010, 2: 328–344
- 30 Meng Q M, Liu H Y, Zhao Z M. Methods, Design, and Statistical Analysis of Psychology and Education (in Chinese). Beijing: Higher Education Press, 2006 [孟庆茂, 刘红云, 赵增梅. 心理与教育研究方法、设计及统计分析. 北京: 高等教育出版社, 2006]
- 31 Triandis H C. Individualism & Collectivism. New York: Westview Press, 1995
- 32 Markus H R, Kitayama S. Culture and the self: Implications for cognition, emotion, and motivation. Psychol Rev, 1991, 98: 224–253
- 33 Spencer-Rodgers J, Boucher H, Mori S, et al. The dialectical self-concept: contradiction, change, and holism in east Asian cultures. Pers Soc Psychol B, 2009, 35: 29–44
- 34 Nisbett R, Peng K, Choi I, et al. Culture and system of thoughts: Holistic versus analytic cognition. Psychol Rev, 2001, 108: 291–310
- 35 Peng K, Nisbett R. Culture, dialectics, and reasoning about contradiction. Am Psychol, 1999, 54: 741–754
- 36 Norenzayan A, Choi I, Peng K. Perception and Cognition. In: Kitayama S, Cohen D, eds. Handbook of Cultural Psychology. New York: Guilford Press, 2007
- 37 Morris M, Peng K. Culture and cause: American and Chinese attributions for social and physical events. J Pers Soc Psychol, 2009, 67: 949–971
- 38 Rodgers J, Williams M, Peng K. Cultural differences in expectation of change and tolerance for contradiction: A decade of empirical research. Pers Soc Psychol Rev, 2010, 14: 296–312
- 39 Ji L, Peng K, Nisbett R. Culture, control and perception of relations in environment. J Pers Soc Psychol, 2001, 78: 943–955
- 40 Li Z. Path of Beauty (in Chinese). Tianjin: Tianjin Academy of Social Science Press, 2001 [李泽厚. 美的历程. 天津: 天津社会科学院出版社, 2001]
- 41 Nisbett R E, Miyamoto Y. The influence of culture: Holistic versus analytic perception. Trends Cogn Sci, 2005, 9: 467–473
- 42 Nisbett R E, Wilson T D. Telling more than we can know: Verbal reports on mental processes. Psychol Rev, 1977, 84: 231–259
- 43 Jung C G. The Archetypes and the Collective Unconscious. Princeton: Princeton University Press, 1981
- 44 Hazan C, Shaver P. Romantic love conceptualized as an attachment process. J Pers Soc Psychol, 1987, 52: 511–524
- 45 Wagerman S A, Funder D C. Situations. In: Corr P J, Matthews G, eds. Cambridge Handbook of Personality. Cambridge: Cambridge University Press, 2009. 27–42
- 46 Comte A. A General View of Positivism. London: Trübner and Company, 1880 [孔德. 论实证精神. 北京: 商务印书馆, 1996]
- 47 Peng K P, Yu F, Bai Y. Experimental ethics: New challenges and contributions to the understanding of human moral behaviors (in Chinese). Soc Sci China, 2011, 6: 15–25 [彭凯平, 喻丰, 柏阳. 实验伦理学: 研究、挑战与贡献. 中国社会科学, 2011, 6: 15–25]
- 48 Yu F, Peng K P, Han T T, et al. Personality and social psychological view on ethical virtues: The meanings, puzzles and solutions of moral traits (in Chinese). J Tsinghua Univ (Philosophy and Social Sciences), 2012, 4: 128–139 [喻丰, 彭凯平, 韩婷婷, 等. 伦理美德的社会及人格心理学分析: 道德特质的意义、困惑及解析. 清华大学学报(哲学社会科学版), 2012, 4: 128–139]
- 49 Sherman R A, Nave C S, Funder D A. Situational similarity and personality predict behavioral consistency. J Pers Soc Psychol, 2012, 99: 330–343
- 50 Terracciano A, Abdel-Khalek A M, Adam A, et al. National character does not reflect mean personality trait levels in 49 cultures. Science, 2006, 310: 96–100
- 51 Henrich J, Heine S J, Norenzayan A. Most people are not WEIRD. Nature, 2010, 466: 29
- 52 Arnett J. The neglected 95%: Why American psychology needs to become less American. Am Psychol, 2008, 63: 602–614

- 53 Henrich J, Heine S J, Norenzayan A. The weirdest people in the world? *Behav Brain Sci*, 2011, 33: 61–135
- 54 Rothstein H R, Sutton A J, Borenstein M. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Hoboken: Wiley, 2006
- 55 Kahneman K. A proposal to deal with questions about priming effects (Letter flow to Nature), 2012. [http://www.nature.com/polopoly\\_fs/7.6716.1349271308!/supinfoFile/Kahneman%20Letter.pdf](http://www.nature.com/polopoly_fs/7.6716.1349271308!/supinfoFile/Kahneman%20Letter.pdf)
- 56 Zhong C B, Lake V B, Gino F. A good lamp is the best police: Darkness increases dishonesty and self-interested behavior. *Psychol Sci*, 2010, 21: 311–314
- 57 Emmons R A, McCullough M E. Counting blessings versus burdens: An experimental investigation of gratitude and subjective well-being in Daily Life. *J Pers Soc Psychol*, 2003, 84: 377–389
- 58 Linden G, Smith B, York J. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Comput*, 2003, 7: 76–80
- 59 Mill J S. *A System of Logic*. Honolulu, Hawaii: University Press of the Pacific, 2002
- 60 Hume D. *A Treatise of Human Nature*. Oxford: Claderon Press, 1843 [休谟. 人性论. 郑之骧, 关文运, 译. 北京: 商务印书馆, 1980]
- 61 Bushman B J, Anderson C A. Media violence and the American public: Scientific facts versus media misinformation. *Am Psychol*, 2001, 56: 477–489

## Big data and psychology in China

YU Feng<sup>1,2</sup>, PENG KaiPing<sup>2,3</sup> & ZHENG XianJun<sup>2</sup>

<sup>1</sup> *School of Economics and Management, Tsinghua University, Beijing 100084, China;*

<sup>2</sup> *Department of Psychology, Tsinghua University, Beijing 100084, China;*

<sup>3</sup> *Department of Psychology, University of California, Berkeley CA94720, USA*

Western psychology has essentially become an independent discipline, relying on empirical laboratory research. However, the era of big data will dramatically influence the research paradigm of scientific psychology. The characteristics of big data methods coincide with Chinese folk epistemology, which provides Chinese psychology the opportunity to weaken the influence of Westernization and restructure our own psychological systems. We proposed an interactive system of TIAN, DI, and REN as a new Chinese psychological system. Based on this system, Chinese psychology can now study a wide range of subjects, from the cultural and spiritual mental representations to the influence of social situations, and further the individual's psycho-physical experiences of cognition, emotion, reasoning, morality, and behavior. Following this path, Chinese psychology can become, as termed by Auguste Comte, the "the last science". Under this new system, Chinese psychology will build a new paradigm and go beyond the dialectical dilemmas in traditional psychology, such as sample vs. population, individuals vs. patterns, situations vs. experiments, and data vs. behaviors.

**Chinese psychology, big data, systems of discipline, culture, epistemology**

doi: 10.1360/N972014-00977