

# 脑机接口与知行合一新形态<sup>〔\*〕</sup>

肖 峰

(上海大学 马克思主义学院, 上海 200444)

〔摘 要〕作为一个哲学命题的“知行合一”有其专门的含义, 哲学家们也探寻了种种通达这一目标的人文途径。脑机接口的研发和使用为知行合一的达成提供了一种不以身体为中介的新技术手段, 这就是通过脑机接口介导的“以想行事”(“运动想象”与“实现行动”的技术性一体化)来实现“我思即我行”。人脑与脑机接口联结而成的新型的知行合一系统为我们重新理解“行动”“行动主体”等概念提出了新的哲学需要, 也对知行交互中的控制问题、错位问题、归责问题和功能极化问题等提供了新的哲学启示。

〔关键词〕脑机接口; 知行合一; 以想行事; 我知即我行; 行为体

DOI: 10. 3969/j. issn. 1002 - 1698. 2022. 08. 007

知行关系是一个重要的哲学问题, 知行合一被许多哲学家视为在这一关系上能够通达的最理想目标, 这些哲学家还为我们趋向这一目标提供了种种精致的理论分析和人文手段, 这些手段包括道德自律、德性养育和日常修行等。今天, 脑机接口正在为我们开辟一条新的路径, 即通过技术手段来为知行合一赋予新的含义与价值, 开辟新的通达之道。

## 一、知行合一: 从人文形态到技术形态

我们知道, 王阳明所提出的“知行合一”, 首要的目的是作为一种道德理想来倡导的, 其中的

“知”主要是指作为道德意识的思想意念, “行”主要是指作为道德践履的实际行动, 而知行合一所倡导的是道德意识和道德行为的互不分离, 即“良知”与“良行”的统一, 如表现为“见父自然知孝, 见兄自然知悌, 见孺子入井自然知恻隐”。孟子的“致良知, 达良能”, 朱熹的“博学之, 审问之, 慎思之, 明辨之, 笃行之”, 以及“穷理以致其知, 反躬以践其实”也表达了同样的精神实质。这种人文传统的知行合一观, 还被称之为“价值论的知行合一”, 旨在克服不善之念、抵御外物诱惑, 做到思无邪、致良行, 并反对道德教育中的知行脱节尤其是知而不行, 倡导言行一致、守信笃

作者简介: 肖峰(1956—), 《学术界》本期封面人物, 重庆人, 法学博士, 上海大学马克思主义学院特聘教授、博士生导师, 华南理工大学、江西师范大学和中国青年政治学院教授。主要研究方向为科学技术哲学和马克思主义哲学, 出版学术专著《科学精神与人文精神》《信息主义: 从社会观到世界观》《信息技术哲学》等 20 余部, 在《中国社会科学》《哲学研究》《马克思主义研究》等刊物上发表学术论文 320 余篇, 其中被《新华文摘》《中国社会科学文摘》《高等学校文科学术文摘》和《人大复印报刊资料》等转载或摘转近百篇次。

〔\*〕本文系国家社会科学基金项目“脑机接口的哲学研究”(20BZX027)的阶段性成果。

行的修身原则。这一价值论或伦理观的知行合一,也衍生为重要的认识论原则,以此揭示人的认知活动与行动践履之间的互在互渗,如王阳明用“好好色,恶恶臭”为例表明了日常生活中普遍存在的这种意念与行为同时发生的知行关联,这样的知行关联也被引申为一种认识论上的方法论原则,以求克服知行脱节、知行相悖、偏重一端而忽视另一端等知行不合现象。知行合一的这些内涵,主要指向的是道德自律、品德修养、合理的认知和行为方式等,旨在“教育”和“引导”人具有正确的修为和素养,所以可以谓之“人文性知行合一”,这也折射出倡导知行合一,是中华文明的一种悠久的人文传统。

在承认“知行合一”上述原意的基础上,如果引入技术的视角,在今天还可以对其赋予“新解”,这就是通过脑机接口(BCI;Brain-Computer Interface)介导的“技术性知行合一”来开辟实现这一目标的新途径。

脑机接口是这样一种技术,它通过采集和解码人脑进行认知活动时产生的脑信号,将其通过计算机及其算法转化为可以控制人脑外部设备运动的指令,从而形成符合人脑意图的替代身体动作的行动。这样的脑机接口也被称为“面向运动的脑机接口”,其功能就是帮助失去行动能力的残障人士恢复行动的能力。在行使肢体功能的脑机接口中,人凭思想控制设备(从意念控制外骨骼设备到控制假肢)做出动作与环境互动,向对象施加物理性作用,造成“改变世界”的效果,这虽非传统意义上的行动,但确实具有了行动的功能,由此可将其称为“人工行动”。

从认识论上看,知行合一中的“知”,是行动的意念、动机、意图,是“想要做什么”的认知状态。这样的认知状态,在通常的情况下,如果不启动身体的动作,就还不是行,还只是停留在认知阶段。也就是说,当我们谈论知行合一时,是建立在知和行“合而不同”的基础之上的,即行动或行为具有与认知不同的含义,它是一种能够改变客观对象的活动,或马克思所说的“感性的对

象性活动”,即外在可感的身体活动,是与脑内无法直感的思维认知活动不同的另一种活动。日常所说的“思想的巨人,行动的矮子”,就表明行动是与思想(知)不同的活动;毛泽东在《实践论》中将认识和实践的关系简称为“知和行”的关系,则更直接地表明了“行”(行动、行为)是与“知”(认识、思想)结成对立统一关系的哲学范畴。

基于行动与认知的这种功能性区别,在身体不动的情况下就不可能取得行动的效果。因为对于正常的人来说,行动都必须通过身体的动作才能实现,甚至行动哲学对“行动”的界定就是“有意图的身体动作”,其言外之意就是:不动手、不践行的“想”“思”或“知”,就是“空想”“纯思”或“罔知”,根本不属于真实的行。而使用脑机接口后,人脑中属于认知范畴的意图或运动想象就可以“自动”化为BCI系统的人工行动,与行动相关的“想”“思”“知”在此时都具有了身体行动的效果。迈克尔·杨(Michael Young)将这种效果描述为“一动不动、无所作为的身体,现在可以极大地影响世界。迄今为止,人类影响世界的任何手段都需要移动身体的外壳,即使只是轻微移动,如说话或眨眼。BCI首次使人类能够在没有身体作为中介的情况下‘通过思想’去做事情”,“BCI第一次能够主动(或因果地)在不移动人体的情况下实现世界变化”。<sup>[1]</sup>可见,脑机接口提供了一种新手段来实现知行合一中的行动,造就出技术手段介导的知与行的同步发生,即BCI加持下的“一念发动处便是行”,这也是另一种形态的知行合一,或称“技术版的知行合一”。

第一,脑机接口所介导的是基于“以想行事”的知行合一。

通常意义上的行是由人的知(意念)引发、支配与控制的人的身体性的感性物质活动,其交互过程是由意识活动引起身体活动,身体活动引起外物运动,这是知行传递的一般模式:脑中的意念(知)一定要启动肢体的物理运动才会成“行”,脑的内部心灵过程一定到转变为身体的外部动作才会现“行”。这就是以身体为中介的

知行转换,身体在这里充当了知行之间的接口,只有通过身体才能表达和实现作为“知”的行动意图与愿望。而当脑机接口行使知行联结的功能后,人的想法可以直接转变为人工行动,使得“在不需要动手的情况下,一个人想什么就是在做什么”,<sup>[2]</sup>尤其是当脑机接口对脑信号的处理足够迅速时,就可以做到一旦脑中有“打算做”的想法,脑机接口系统就能同步响应而使其“做成了”,此时甚至“依靠身体动作或言语来表达意图已经变得毫无意义。你的想法会被脑机接口有效而完美地转化为纳米工具的细微操作或者尖端机器人的复杂动作”。<sup>[3]</sup>

人脑在联结面向运动功能的脑机接口后,行动在脑中的预演(准备行动的认知活动)和在体外的实施(真实行动)被集合为同一过程,一定程度上达到了知就是行、知时就行、即知即行的浑然一体。在传统的认识论视野中,仅停留在头脑中“想做某事”时并不算是真正在做某事,还不能说是进入了行动的状态;而连接了面向运动的脑机接口后,作为知行联结装置的脑机接口使得人脑中的意念与体外的人工运动系统形成对接和贯通,这就是知和行之间的技术性一体化。这无疑是一种新型的知行关联形态,它改变了传统的知行联结方式,使得基于身体作为知行接口时需要“以手行事”(“亲手”做事)或“以言行事”(指使别人做事或“告诉机器做事”)变为施泰纳特(Steffen Steinert)等人所说的“以想行事”(doing things with thoughts),它具有“非具身”(disembodied)的性质,不需要涉及身体或肌肉的运动,<sup>[4]</sup>“仅通过大脑活动就将思想转化为行动”,<sup>[5]</sup>达到了“我思即我行”(Cogito is Facio)的状态,这就是通过脑机接口来支配智能机器为我们去做我们想做的事,所以这是一种基于“以想行事”意义上的知行合一。

第二,脑机接口所介导的是基于人机分工与人机融合的知行合一。

由脑机接口介导的新型的知行合一,还是一种基于人机交互的知行贯通,它是在知行间新型

分工的基础上形成的,这就是人机之间的知行分工:人自身不再充当行的工具,从而完全从工具的地位中解放出来;在将来,当脑机接口不仅是治疗的手段,而且是普遍的劳动工具和增强手段时,人机之间的这种知行分工将更为泛在。在这样的背景下,一方面是机器对人的身体之体力行动功能的“替代”,另一方面正是在这种替代中又实现了新型的知行合一:一种“有分才有合”意义上的知行合一,即“人想”与“机器做”的“人机合一”,此即基于人机协作的“心想事成”,<sup>[6]</sup>而且是更高水平的心事成。因为通常的“心想”和“事成”之间总有一定的距离,从心中意图的产生到将其变为现实需要一个漫长的过程。但一旦接入(高水平的达到“透明”或“上手”状态的)脑机接口系统后,这个距离被电子传输的通道所缩短甚至消除,作为意图之“心想”和作为结果之“事成”被融合为“一体化”的过程,知的同时就在行,意念中“打算做”“尝试做”与人工系统的“实际做”通过脑机接口实现了同步化,“脑动”的同时启动了外在的“物动”。

这种基于人机分工基础上的知行合一水平还在不断提升。脑机接口技术是不断发展的,从植入单电极来获得少数神经元的脑信号,到植入多电极来获得大面积神经元信号,从充当传感器的电子芯片与脑组织的有创对接,到生物芯片与神经元的无缝对接,都是已经出现或将要发展的趋向。在这样的进程中,使用者从人机合一中所体验的“知行合一感”也会不断提升。目前的脑机接口技术水平还不高,所以使用者形成基于人机合一感的知行合一感的程度往往不同。在一项相关的研究中,研究人员向使用者征询如下的问题:“在使用BCI设备时,你能直接把自己的注意力集中在你想要做的工作上吗?换句话说:你能忘记技术只做你想做的事情吗?”20个受试者中有19人回答:是的,经过大量的训练就能够做到(即具有知行合一感)。其中有6人进一步回答:通常是,它是没问题的。还有的甚至达到了高度的人机合一状态,即较高度度的知行合

一感,“是的,当我看到假手的开合时,我感觉就像是我真实手的自然运动。”当然也有的未获得知行合一感,“不,实际上工具在培训中起着核心作用,忘记它只关注手上的事是不容易的,这对我来说甚至是不可能的。”<sup>[7]</sup>这就表明了脑机接口的当前使用可以造就出不同的人机合一状态即知行合一感,可以说,当BCI成为上手的透明工具或进入到“人机合一”状态后,人脑中的“知”就会感觉不到其存在,BCI系统被使用者有机整合、同化成了自己的一部分,他们在使用脑机接口控制的假肢时,如同是无意识使用自己的肢体在有效地控制环境,专注于目标而“忘记”了机器设备的存在,此时他们关注的不是“看看我的手臂做了什么”,而是“看看我做了什么”;不是“我向右移动计算机”,而是“我向右移动”,由此形成了种种基于人机融合的知行合一的神奇体验。<sup>[8]</sup>

第三,脑机接口所介导的是基于“接口”之强大联结功能的知行合一。

脑机接口技术之所以能贯通知和行,也在于它所具有的“接口”性质,以及它既是认知工具(能读取人的思想、意图),也是行动工具(能将读取的脑信号编码为外部机器可以理解的信息指令从而导致受控的人工行动)的双重属性,从而能够成为集合两者功能的装置。面向运动的脑机接口,一端联结的是脑中的认知活动,另一端联结的是外周设备的人工行动,其功能是将人脑中的行动意图传递到能够施动的技术系统中,使工具与对象之间发生实际的交互,产生改变外部对象的行动效果。在这个过程中,人脑内的“知”与人体外的“行”由脑机接口联结起来,使得脑机接口也成为功能意义上的“知行接口”。在这个具有接口功能的集合体上,实现着人的认知和人工行动之间的对接、传递和贯通。脑机接口所创造的知行融合效应表明,人类追求“接口技术”的努力是持续不断的,通过新的接口技术将不同的现象或方面贯通起来,在新的联结之处造就出人所需要的新功能。脑机接口作为一种

新的接口技术,使得人类的知行贯通能力得到极大的增强,消除知行隔离或脱节的水平也不断提高,新形态的知行合一的威力也不断彰显。

脑机接口所介导的技术性知行合一与传统的人文性知行合一具有不同的特点。人文性知行合一主要是为了在伦理观或知行观的偏好上进行“补偏救弊”:“有人偏于冥行,便教之知以救其弊;有人偏于妄想,便教之行以救其弊。必使他达到明觉精察之行,真切笃实之知,或知行合一而后已。这样一来,知行合一便成了理想,便须努力方可达到或实现的任务。”<sup>[9]</sup>而技术性知行合一主要解决的是“生理能力”或“自然功能”上的缺失问题,即行为上的失能,这是一种特殊的“知行不合”,一种受制于身体病残的极端化的“知易行难”,或“知能行障”,还包括在能力修复基础上的未来增强——突破人的生物学限制的功能增强,使目前或自然状态下“力不能及”的认知意图能够借助增强系统兑现为现实的行动,所解决的是生理问题、技术赋能问题,而非伦理教诲或人文素养问题。

人文性知行合一所侧重的是对某种理念重要性的“阐释”,主要是在“论”的基础上展现出来的一种人文理想或追求,也是在“应该”意义上的价值规训或修为倡导。而基于脑机接口的知行合一,是经过技术努力后在客观上可造就的知行合一,它偏向于具体的技术操作,侧重于在“做”中实现行的实在效果,所形成的是技术性的成就,它使得知行合一不再仅仅是一个道德诉求或哲学思辨的命题,而且成为向科学技术开放的目标,成为可由技术手段在功能上实现的事实,一种具有实在性的客观现象,从而也是对价值性、人文性知行合一的一种重要补充,它赋予了作为哲学命题的知行合一新的含义,也将人类的知行合一能力提升到了一个新的水平,推进到了一个新的空间。

## 二、新型的知行合一系统与新型的行动主体

脑机接口介导的人与机之间的融合达致知



与行之间的贯通,造就了一个技术性的新型知行合一系统,这个系统由人和 BCI 两大部分组成,这两大部分具有如上所述的人机之间的分工,在这种分工中,当“行”的功能由 BCI 系统来承担时,行动的主体又是谁?人还是这个新型知行合一系统中的行动主体吗?这是技术性知行合一“对“行动”和“行动主体”的含义所提出的问题。

其一,新型的知行合一系统扩展了行动主体和行动的外延。

通常来说,行动主体是具有行动能力的人,是能行动的人,是具有身体动作能力的人。因为行动就是“意愿 + 一系列适当的身体动作”,两种能力缺一不可。如果这样来看,那些丧失身体动作能力的人(即那些“基本完整的思想被困在几乎完全瘫痪的身体里”<sup>[10]</sup>的残障人士),其知和行在他们那里处于一种能力上的脱节或不连续状态,他们作为主体的功能只能止于认知而不能接续为行动,因此不能称其为行动主体。但在脑机接口出现后,情况发生了根本性变化,因为“脑机接口提供了一个无障碍的渠道,可将个人某些行动的目的传达给外部设备,例如计算机,语音合成器,辅助设备和神经假体。所以脑机接口系统特别适合患有严重运动障碍的人”,<sup>[11]</sup>这样的系统可以使他们“通过大脑深处集中的思想力量”来实现自己与外界打交道的目标,<sup>[12]</sup>重拾行动的能力,或重新成为行动的主体。这样一来,新型的知行合一系统,既使“行动”的概念具有了更大包容性,它“涵盖一些(但不是全部)由 BCI 介导的事件”;<sup>[13]</sup>也使“行动主体”随之扩容,使那些不具有身体行动能力但借助脑机接口可以脑控人工行动的人也成了现实的行动主体。

其二,新型的知行合一系统造就了“延展的行动主体”。

在脑机接口造就的知行融合系统中,知和行的能力体现在整个人—机系统中,此时,当人只承担“运动想象”的职能时,他似乎不再是完全意义上的行动主体。换句话说,当行动对象的改变是由人和脑机接口共同带来时,此时的行动主

体似乎很难归结为单独的一方,这就引出了经常出现在人机系统中的“担责”问题:人脑借助脑机接口所产生的行为后果(尤其是不良后果)由谁来负责?弗莱克(Rutger Vlek)等人列举了这样一个事例:想象一个用户——我们称他为弗兰克——试图在 BCI 的帮助下执行操作,弗兰克正在想象他的左手在移动,以便让机械手端起一杯热咖啡,这样坐在附近的人(路易斯)就能拿到它。但在这样做时出了问题:咖啡溅到了路易斯身上,使他的衣服被弄脏,皮肤也受到轻微灼伤。经调查发现,弗兰克的脑控活动并无过错,是 BCI 的某些环节没有正常工作导致了上述结果,即弗兰克的大脑状态在这个不幸的结果中没有起因果作用,但弗兰克还是为此感到内疚,并对此付出了赔偿。<sup>[14]</sup>这表明人与脑机接口结合后“谁是行为的主体”进而“谁是行为后果的责任人”并非一个简单的问题,很难简单地归结为人或 BCI 机器哪一方。

从脑机接口的工作机理上,德鲁(Liam Drew)对这种新型的知行合一系统中的归责问题进行了分析:脑机接口在使来自大脑的信号到达假肢之前,需要被人工智能软件处理,如用机器学习软件来对神经活动加以分析和解码,而机器学习软件是通过生成无法预测、难以或不可能理解的算法来学习分析数据的,这就“使整个领域充满了动荡”,相当于在一个人的思想和代表他们行动的技术之间引入了一个未知的,也许是不可解释的过程。研发人员意识到,当某些计算留给 BCI 设备,由这些设备来预测用户下一步将要做时,假肢的工作会更有效率。例如看似简单的动作,比如拿起一杯咖啡,实际上非常复杂:人们会下意识地执行许多计算。在假肢上安装传感器和自动产生连贯动作的机械装置,可以让用户更容易地执行任务。但这也意味着,很多机械臂所做的事情实际上并不是由用户指挥的。BCI 中的智能算法从先前的数据中学习,并指导系统根据过去所做的事情作出决策。但是,如果算法不断提示用户的下一个动作或单词,而用户

所做的仅仅是批准该选项而已,那么该动作或该单词所代表的消息究竟是由谁发出的,将变得模棱两可,<sup>[15]</sup>由此出现责任归属的复杂性,这无疑表明了究竟谁是行动主体的复杂性。

如果说第一种情形已表明了行动主体含义的扩展,使得在自然状态下不具备行动主体功能的人在技术介入下成了新的行动主体,那么当脑机接口本身也具有为行动后果担责或部分担责的地位后,它也在某种意义上可视为行动主体的一部分,从而使行动主体的含义进一步被扩展,脑机接口此时从“延展行动主体”的意义上为行动主体进行了再度扩容。

其三,新型知行合一系统中的主体感问题。

行动是把想象变为现实,当脑机接口也能将人的想象变为现实时,或者说对于那些不能依靠自己的身体行动来将想象变为现实的人来说,如果移除脑机接口系统,他们就不能成为行动主体,因此脑机接口是这类人群能否重获行动能力的决定性因素。脑机接口的这种关键作用无疑也会使用户产生主体身份的质疑:自己是真正的行动主体吗?用迈克尔·杨的话来说:他们在BCI介导的事件中是否能体验到一种主体感,以便将这些事件归为自己的行为?<sup>[16]</sup>这一问题生动地体现在一些脑机接口使用者的心理感受中,从而引发了他们的“主体感”问题:使用脑机接口“做事”或行动时,他们常常会提出“这件事是我做的吗?”“我是在行动吗?”“我是行动的发起者和承担者吗?”之类的疑问,即“我”使用脑机接口所做之事是否为“属我”的行动?从对相关用户的了解来看,“如果他们没有施动者的感觉,就可能不会将脑机接口介导的运动归因于自己,由此就会缺少行动的感觉,并将脑机接口介导的事件归为不受自己支配的一类现象,尽管他们实际上确实启动或控制了事件的进行。此时他们并不认为自己就是施动者。”<sup>[17]</sup>就是说,即使旁观者认为他们是行为的主体,脑机接口的使用者自己也可能没有作为行动主体的感受,其根源在于人脑中正在想什么与脑机接口系统正在做什

么之间可能因种种原因发生了错位而未能协调一致,造成知行不能合一的感受。这也表明,面向运动的脑机接口如果还不能达到知行融合的水平,在脑中的意愿与人工系统的运动之间未形成有机的融贯,从而达不到技术性的“心想事成”之境界时,用户就会在使用脑机接口时缺乏主体感。从这个意义上,使用者是否具有主体感(也是一种知行合一感),是脑机接口是否达到技术性知行合一的判别标准,也是体现脑机接口水平的重要标志。

其四,新型知行合一系统中主体边界的确认。

如果把脑机接口所造就的延展系统也纳入行动主体的范围内,成为其中的一部分,那么行动主体的边界在哪里?将植入式脑机接口的植入部分视为主体的一部分似乎在“空间”上比较容易接受,但那些未植入脑中却又与人脑在外部空间上相连接的脑机接口器件是否为行动主体的一部分?自然的肢体是行动主体的一部分,这是毫无疑问的;而延展的人工肢体是行动主体的一部分吗?若是,则这种肢体如果可以无限延长,那么其边界在哪里?若不是,那么对于没有联结人工肢体的残障人士来说就不具备行动能力,进而也难以称其为行动主体,这又与前面所说的脑机接口使行动主体的含义得到扩展相矛盾。

目前的解决思路,一是对于人是主体的看法毫不动摇,将一切技术性的东西都视为工具、手段,即使深度融入人脑中的脑机接口也不例外,此时人作为行动主体的地位并不受到发展的任何影响,这也可以从人机之间、知行之间的“主从关系”上得到合理的解释。在基于人机分工的知行分工中,由于人贡献知,BCI系统贡献行,而两者之间具有知为主、行为从的关系,由此也决定了人为主、BCI为从(人主机从)的关系,所以人是整个知行过程的主体,从而归根结底也是行动(包括人工行动)的主体。这种“知主行从”的关系也是王阳明的知行合一观所主张的,他认为知是行的主意,行是知的功夫。知是行之始,行是知之成。贺麟对此进一步解释道:知是

行的本质,知决定了行的意义,使行由一种“物理行为”升华为一种“意义行为”;知决定行还表现为不能做他不知之事;另外,行的质量也是由知所决定的,而且知是目的,行是实现目的的工具。<sup>[18]</sup>所以,当脑机接口介导下知和行分别由人和机器来承担时,人所充当的“知”的主体职能充分保证了他作为整个知行过程主体的地位,从而也保证了人从根本上仍旧是行动主体的地位,因为此时的人工行动是人所掌控的、符合其意愿的行动,而非机器的盲目运动。这样,脑机接口使得知为主行为从的关系得到了技术化的验证,同时也确保了新型知行合一系统中人所具有的行动主体的地位。

第二种解决思路,是在人作为主体的初始含义上附加新的因素,类似“人—机主体”所表达的含义。“人—机主体”是讨论人工智能介入、辅助和增强人的能力时所形成的概念,这种人—机主体不仅有认知的功能,而且还有行动的功能,从而产生了可称之为“人—机合作的行动”现象。类似地我们可以将基于脑机接口的行动主体视为“人脑+脑机接口+人工肢体”组成的行动主体,或以“延展主体”来表达。这里的“延展”表明了它与原本的以人的身体为边界的主体已有所不同,可以理解为行动主体存在的新形式,也是BCI技术赋能于主体达到了一定程度后必然出现的结果。从实际运行的层面,还有人将这一视角表述为主体职能的“共同分担”。如对于配有智能控制技术的脑控轮椅:“在这样的控制方案中,责任是由用户和系统共同分担的,用户给出高级命令,系统执行低级交互,具有或多或少自治程度。利用这种所谓的共享控制原理,研究人员已经证明了通过无创脑机接口用意念控制复杂移动设备的可行性”。<sup>[19]</sup>在我们看来,即使脑机接口被纳入行动主体的范畴,也不是它的所有部分都可称之为行动主体的一部分;而且即使可称为行动主体的那部分,也具有既是主体也是工具的双重属性。

第三种解决思路,是用一个更大的囊括性更

强的概念 agent 来指称脑机接口、人工智能等介入后的行动主体。目前 agent 的汉语表达有“主体”“代理”“智能体”“智能主体”“自主体”“自为体”“中介”等词汇,尤其是还有“施动者”“行为体”的意思,这就与“行动者”具有较大的吻合性,此种意义上 agent 不再区分是人还是其他施动者,也不再区分它是硬件还是软件,是碳基生命还是硅基技术,只要能施加行动的某种实体都称为 agent。“在这种情况下,我们需要接受的是,在先进技术中,人类和非人类 agent 之间的交织使得各种 agent 的任务变得难以区分。因此,将整体产出的责任归于单一 agent 似乎是有问题的。”<sup>[20]</sup>这样来看,目前的脑机接口具有 agent 的部分属性,尤其是“代理”的属性,将来还可能具有“自为体”的属性,从而成为一个完全意义上的 agent。当然,人早就是充分且完全意义上的 agent。在这个意义上,agent 可以在一定程度上表达新型知行合一系统中所涉及的动作主体方面的属性,或者说,脑机接口部分在新型的知行合一系统中是否具有行动主体或行动者的属性与地位,可以在更大的 agent 概念的扩充和统一中获得解释。

### 三、知行合一新形态与哲学新问题

脑机接口所造就的技术性知行合一,除了引发上面专门探讨的行动主体问题外,还带给我们更多的哲学新问题。

第一,知行交互中的有效控制问题,尤其是基于知行区分的知行合一问题。

在脑机融合所造就的知行合一新形态中,我们对“行动”的传统理解发生了改变。在脑机分离的时代,身体是从意图到行动的唯一通道,行动必须是具身的现象;而在脑机融合的时代,行动的意图可以迈过身体去实现,即在身体不动的情况下取得行动的成果,达成“我思即我行”的知行联结。但这样的我思即我行也会带来新的问题:做与不做难以区分。在联结上高效的面向运动的脑机接口后,如果只是“想一想”而先不



去行动的话,就难以确保可以停留在这个阶段,因为高度灵敏的脑机接口只要一捕捉到人脑的“想法”,就可能同步将其付诸机器的运动,即使对于该想法的持有者来说,还没有将其付诸实施的打算,这就是“我思即我行”有违人意的情形。就是说,人即使只想停留在“动作想象”而非“动作执行”的阶段,也可能被脑机接口推向动作执行的阶段。那么在联结脑机接口的情况下如何才能保持只停留在运动想象、行为预演、实践构思的阶段,难道只有拔掉脑机接口的插头才能做到吗?一些研发者发现,脑机接口执行“停止行动”的指令可能比执行“开始某种行动”的指令更加困难,以至于在脑控轮椅的驾驶中,用户无法将这种轮椅停在空地上,只能将轮椅停靠在障碍物的周围才能确保其稳定地停下来。<sup>[21]</sup>这也从另一个角度表明了,在知行合一的新形态中,在认知活动时想要停止行动的难以实现,也就是将知与行暂时分别开来的难以实现。可以说,达成知行合一的前提是知行的区分,如果不能达成技术性的知行区分,知行合一就是没有区分度的“知行混沌”,就不是我们所期望的有实际意义的知行合一。

第二,知行交互中的匹配与纠错问题。

基于脑机接口的知行合一新形态,需要在形成如前所述的主体感的基础上才能达成,即脑机接口的使用者要能够感受到某行为是由自己的意志引起的,由此才能形成知行合一的感受。此时需要使用者感受到意愿的优先性,即意愿在逻辑关系上先于行动;意愿和行动关联的排他性,即意愿是行动的唯一明显原因;以及意愿与动作的一致性,即思想与行动之间的兼容。弗莱克等人指出,在使用面向行动的脑机接口时,一致性对于使用者具有主体感(知行合一感)有着特殊的意义,此时运动想象被视为驱动脑机接口的心理任务,参与者需要在心理任务和执行器输出之间建立起新的映射,通过学习和训练,不断强化某一心理任务和受控目标运动(如光标移动)之间的相关性,形成两者一致的图像,这样才能增

强他们对自己是某一动作施动者的判断,即获得该行动的主体感,<sup>[22]</sup>亦即知行合一的感觉。

但在使用BCI不熟练时,难免会发生知行错位的情况,如在目睹一件事(认知某一事件)与做一件事(行)之间产生混淆。即使在没有使用BCI的情况下,也会出现两种相反的混淆,一是仅仅是目睹一件事的时候却认为是正在做某事,即有做某事的感觉。如在替代性的控制实验中发现,人们即使知道别人执行了指令,也会对他人的动作具有施动的感觉,这种现象被称为“控制幻觉”。<sup>[23]</sup>二是自己正在做某事,但却认为没有做某事,即并不认为自己是某事的施动者。

使用脑机接口时未达成知行合一的这种知行错位或知行相悖还可能更复杂的情形,比如,为实现良知与良行的合一,就需要避免“恶思”“恶行”的发生,但人脑中由于种种原因难免会出现认知偏差的情况,这就是奥布罗链(Fiachra O'Brochain)等人所指出的:“我们对思想的控制权少于对作为小工具驱动器的身体的控制权。”<sup>[24]</sup>在以身体为中介的知行过渡中,即使发生了认知的偏差,也可以通过身体行为的控制来过滤和阻隔这种偏差。像那些有害的意念意图,可以通过避免“身体的代价”对其加以阻隔和限制,即通过身体的不响应来不使其产生行动效果。身体对脑中行动意念的响应所具有的这种选择性,成为知与行之间的相互“纠偏”或相互控制,这也是以逆向的方式表现出来的知行合一。相比以身体为手段的行动来说,以脑机接口为手段时对于行动效果的控制则会更加困难,因为此时行动的意念不再受到身体的阻隔和限制,只要脑机接口能够将其读取和解码为机器指令,人的体外延展系统就会去“忠实”地加以执行,从而将人的行动意念变为行动结果,无论这种意念是好是坏,是善是恶,是建设性的还是破坏性的。换句话说,对于大脑发送的任何指令(无论是良知还是邪念),当脑机接口“不懂得”趋利避害时,只要它能做得到,它都会被不折不扣地执行。在传统的由知到行的过程中,人尽管难以避



免“胡思乱想”，但身体不会将其都变为行动，所以“行动失序”的情形较少发生，尤其是对于精神总体正常的人来说更是如此。但在脑机接口的介导下，只要 BCI 足够敏感，人脑的任何胡思乱想都可能被脑机接口所接收，从而导致外部设备的无序甚至“疯狂”运转，形成不可思议的后果，脑机接口由此有可能“降低了我们控制对象乃至我们自身的能力”。<sup>[25]</sup>这也表明，脑机接口所造就的技术性知行合一，还存在着什么样的知与什么样的行相匹配的问题，尤其是为了实现良知与良行的合一，需要有对知和行进行“查错”“纠偏”的机制，这也正是人文性或价值性的知行合一需要植入到技术性知行合一的地方。

### 第三，人机交互中的行动归责问题。

上一节探讨新型知行合一系统中人和机器谁是行动的主体时，已涉及行动的归责问题，这里主要分析潜意识通过 BCI 所转化的行动以及脑云接口介导的行动之归责问题。

人的心智活动中，除了有能被自己觉察的意识活动外，还有大量不被觉察的潜意识活动，人们即使在清醒状态下也无法自觉地控制这些潜意识。在使用脑机接口时，潜意识有可能对其正常工作产生意外干扰。当潜意识不受控制地传递给脑机接口并转换为行动时，会导致种种意愿之外的实践后果。这就是“不自觉”的“知”对于行的错位匹配，此时脑机接口系统的人工行动不是来自人脑中自觉的运动想象，而是来自其不自觉的潜意识，BCI 由此成为“潜意识的执行器”。<sup>[26]</sup>奥布罗链等人特别分析了潜意识支配脑机接口的行动时可能会出现复杂情形：在脑机接口采集脑信息的灵敏性达到一定程度时，将有能力探测到人脑的潜意识活动。如果由于 BCI 对其用户的潜意识进行了响应而发生了事故，那么显然存在责任如何分配的问题。此时要使用者来承担责任是很困难的，因为他们既不知道自己在想什么，也无法控制触发某些动作的潜意识。这种 BCI 可以利用大脑潜意识活动的能力会从根本上使我们的主体概念复杂化，<sup>[27]</sup>从而

也使归责的问题复杂化，尤其是当潜意识通过知行接口引发了破坏性的后果时，我们会面对更难处理的问题。这时，我们既要提防脑机接口被“不怀好意”的人用来“做坏事”，也要避免没有恶意的“好人”因为联结了脑机接口而在潜意识作用下的“无意为恶”，这也可以看作是阿伦特 (Hannah Arendt) 所说的“平庸之恶”的技术版本。凡此种种，都是良知与良行未能实现合一的“知行相悖”现象。

从理论上可以认为，一个人要为某事担责，一个重要条件是需要他意识到自己正在以一种特定的方式行事。而潜意识支配的行为通常被认为是非自愿的，所以至少其不会因此负全部责任，但由于脑机接口这种特殊的知行联结所具有的破坏性后果的风险存在，所以必须寻求减少甚至避免这种风险的路径。比如可以从技术上努力，将脑机接口提升到能够自动识别和排除人脑中产生的不合理的行动意图，使其具有纠错、纠乱、过滤的功能，能够识别脑信号意外的和无意的神经激活，从而智能化地阻止某些不合理的危险的“知”借助脑机接口而走向“行”。也就是说，基于脑机接口的知行联结除了要确保行为效果的合目的性，还必须解决种种“知行失配”（即未能达成良知与良行的合一）的问题，使得合目的的“行”必须联结的是自觉、理性、有序以及善良的“知”，通过各种手段（包括技术手段）来避免人机之间的知行错位。相信人工智能水平的提高将有助于解决这一问题。

此外，当脑机接口发展到更高水平，还会产生人脑与电脑云的联结，即所谓“脑—云接口”，此时众人的“知”可以结成一个“认知共同体”来控制应用系统的“行”，结成一个更加庞大的知行合一系统，这个系统中众人的认知融合所导致的可能是在众人的意识交互状态下进行决策、实施行动，即“由 BCI 连接的群体作出集体决策和采取集体行动。在这样一个全新的场景中，可能无法确定哪个人对哪种思想负责，并最终为行动负责。以这种方式使用的 BCI 不仅会挑战我们

的责任观念,还会挑战个体和代理主体的观念。假设BCI确实使人们能够连接在一起执行任务,将对用户的责任分配产生重要影响”。<sup>[28]</sup>尤其是当融合心灵控制的行动所形成的后果如果是意外的破坏性的,就有可能很难准确溯因,找到究竟是谁的“知”主导了如此的“行”。对此,有学者倡导还需要有集体责任的概念,<sup>[29]</sup>需要对不同维度的责任进行全面的分析和整合,来为这个新型的知行合一系统中的责任追溯问题搭建一个复杂而完整的治理框架。

#### 第四,知行交互中的主体功能极化问题。

脑机接口不仅可用于恢复残障人士的行动能力,而且可用于为正常人做事,逐渐替代人的一切体能性行动,但在这个过程中也包含着知行分裂的新可能:“行”的功能与人渐行渐远,所有基于体能之劳作都移交给BCI系统全权代理,人以纯粹“动脑”的主体代替了过去需要“动手”所充当的工具。当脑机接口极大地增强了“替代性动手”的效果时,是否意味着人自身可以是纯粹的“知者”而不再是“行者”?当人自身只需脑中的知就能达到一切目的时,是否也意味着在人身上了发生了知对于行的消解和侵吞?这显然是脑机接口的功能极化后,导致的人的功能的极化,从人之为人以及人的全面发展的需求来看,我们似乎难以选择这样的极化生存方式,此时人文性知行合一的介入与补充再度显得必不可少。

还需要指出的是,基于脑机接口的知行交互和转化即知行合一,虽然迈过了身体这个物质性的中介,但并不意味着“知”或意念可以直接造就行的效果,即作为“精神”的知在此时也不是凭空就转化为可改变物质的行,而是仍然要借助作为物质和硬件系统的脑机接口及其相应的外周设备,并且还要借用外部的自然力(如电力)才能完成这样的转化。没有人所创造的这些物质系统和借用的外部能量,“知”无论如何神奇也不可能转化为行,在这个意义上,“知”也仍然是不可以恣意妄为的,它能在多大程度上转化为

行,仍然受限于脑机接口和实施器具的技术水平,因此这里所验证的是一种基于唯物论的知行合一,而非一种基于唯心论的知行合一。

#### 注释:

[1][16] Young M., “Brain – computer interface and philosophy of action”, *AJOB Neuroscience*, vol. 11, no. 1, 2020, pp. 4 – 6.

[2][7][10][12][14][19][22][24][25][26][27][28]  
[29] Gerd Gröbler and Elisabeth Hildt (eds.), *Brain – Computer Interfaces in Their Ethical, Social and Cultural Contexts*, Dordrecht: Springer Science + Business Media, 2014, pp. 102, 122, 150, 111, 200, 26, 201, 167, 168, 176, 167, 167 – 168, 163 – 182.

[3][巴西]米格尔·尼科莱利斯:《脑机穿越:脑机接口改变人类未来》,黄珏华、郑悠然译,杭州:浙江人民出版社,2015年,第IX页。

[4][13][17] Steinert S., Bublitz C., Jox R., et al., “Doing Things with Thoughts: Brain – Computer Interfaces and Disembodied Agency”, *Philosophy & Technology*, Vol. 32, 2018, pp. 457 – 482.

[5] Lee J., “Brain – computer interfaces and dualism: a problem of brain, mind, and body, AI & society”, *The journal of human – centered systems and machine intelligence*, vol. 31, no. 1, 2016, pp. 29 – 40.

[6] 肖峰:《知行接口及其哲学分析》,《东北大学学报(社会科学版)》2014年第2期。

[8][20] Kögel J., Jox R. and Friedrich O., “What is it like to use a BCI? — insights from an interview study with brain – computer interface users”, *BMC Medical Ethics*, vol. 21, no. 2, 2020, pp. 1 – 14.

[9][18] 贺麟:《知行合一新论》,载《贺麟选集》,长春:吉林人民出版社,2005年,第385、373 – 391页。

[11] Wolpaw J., Birbaumer N., Heetderks W., et al., “Brain – computer interface technology: a review of the first international meeting”, *Rehabilitation Engineering IEEE Transactions on*, vol. 2, 2000, pp. 222 – 225.

[15] Drew L., Agency and the algorithm, *Nature*, vol. 571, no. 7766, 2019, pp. S19 – S21.

[21] 张瑞:《面向重度残疾人的脑机接口功能辅助研究》,华南理工大学博士论文,2016年,第20页。

[23] Wegner D., Sparrow B., Winerman B., et al., “Vicarious agency: Experiencing control over the movements of others”, *Journal of Personality and Social Psychology*, vol. 86, no. 6, 2004, pp. 838 – 848.

[责任编辑:刘 鏊]