

# 心理信息学：网络信息时代下的心理学新发展

薛 婷<sup>1,2</sup> 陈 浩<sup>3</sup> 赖凯声<sup>3</sup> 董颖红<sup>4</sup> 乐国安<sup>3</sup>

(<sup>1</sup> 天津中医药大学管理学院, 天津 300193) (<sup>2</sup> 阿姆斯特丹自由大学社会科学学院社会学系, 阿姆斯特丹 1081HV, 荷兰) (<sup>3</sup> 南开大学周恩来政府管理学院社会心理学系, 天津 300071)  
(<sup>4</sup> 鲁东大学教育科学学院心理系, 山东烟台 264011)

**摘 要** 在见证了社会计算与计算社会科学的勃兴之后, 心理学和信息科学在网络信息时代相遇, 催生出心理信息学这门新兴交叉学科。该学科旨在利用计算机与信息科学技术, 通过开展网络调查和实验、开发和利用移动设备、建立数据库和分类体系、使用开源软件和数据挖掘技术等, 获取、整理和分析心理学研究资料。其对大规模复杂数据的掌控能力, 使心理学家可以从新的高度和视野, 针对现实和虚拟网络中的个体与群体心理现象展开研究, 重新检视已有理论假设并探索新发现。但是, 心理信息学在展现巨大潜力的同时, 也在技术和研究设计、伦理和研究规范、研究重心与人才培养等诸多方面面临挑战。

**关键词** 心理信息学; 心理学; 信息科学; 计算机技术

**分类号** B849:C93

网络信息时代的到来, 不仅改变了人们的生活方式, 同时也给学术研究范式带来深刻变革。现实世界与网络世界紧密关联, 彼此缠绕, 人类行为痕迹以数据形式大量涌现。心理学研究者愈来愈依赖计算机与信息科学技术, 以探索个体或群体的心理和行为规律。应时代特征和学科发展之需诞生的心理信息学(Psychoinformatics), 正是心理学研究者积极尝试利用计算机和信息科学技术, 获取、整理和分析心理学研究资料的新学科(Yarkoni, 2012)。它已然取得一些令人兴奋的成果, 昭示着这门新兴交叉学科, 将从研究方法和研究内容上给心理学注入新的活力。

## 1 发展背景：心理学与信息科学的相遇

回顾心理学和信息科学各自的发展历史与特点, 不难发现, 两者的相遇并非偶然, 而是两者

在学科目的、内容上有所交叉的集中体现, 同时也是其各自寻求发展的必然趋势。

以信息作为主要研究对象、以信息运动规律为主要研究内容、以信息科学方法论为主要研究方法的信息科学, 其主要研究目标是扩展人的信息功能, 尤其是智力功能(钟义信, 2002)。在信息科学研究者看来, 智能本身并不是什么神秘的东西, 其本质就是一种利用信息的能力。现在科学早已发现, 人类是通过感觉器官、传导神经网络、思维器官和效应器官的协同合作来认识世界和改造世界。信息科学技术则试图在这四种人类信息器官的功能基础上, 进一步发展感测、通信、计算机与智能、控制等技术, 以扩展人的信息功能, 更好地完成信息的获取、传递、加工、再生和施用等功能(Brillouin, 1956)。信息的普遍存在性和信息科学的核心目标, 决定了该学科需要不断地从其它领域汲取养分来丰富自己。这一过程也促使信息科学技术逐渐成为一种典型的通用技术, 深入到系统科学、控制论、认知科学、人工智能等诸多学科领域中(钟义信, 2002)。

近半个世纪以来, 信息科学先后与经济学、生物学、神经科学等学科联姻, 诞生出经济信息学(Ecoinformatics)、生物信息学(Bioinformatics)

收稿日期: 2014-01-27

\* 国家重点基础研究发展计划(973 计划)项目课题(2010CB731404), 国家自然科学基金重大项目(12&ZD218、14ZDA063), 国家自然科学基金重点项目(12ASH006), 教育部人文社会科学研究青年基金项目(13YJC190025、11YJC190004)。

通讯作者: 陈浩, E-mail: hull1919@gmail.com

和神经信息学(Neuroinformatics)等交叉学科。这些领域综合运用信息与计算机科学技术,探讨了与经济活动有关的信息特征及其运动规律;生物信息的获取、处理、储存、分发、分析和解释;神经系统信息的产生、传输、加工规律及编码、存储与提取机理等问题。这些交叉学科在理解和解释各类经济信息现象,阐明核酸和蛋白质序列数据所包含的结构功能,以及大脑不同层次活动的生理学和心理学意义上,取得了许多令人兴奋的成果(邱均平, 2002; 张阳德, 2009; 陈惟昌, 王自强, 陈志华, 安荣姝, 2001)。相比而言,心理信息学的研究成果一直散落在不同学科背景的新近文献之中。

其实,心理学和信息科学对彼此的关注和互动由来已久。早在 Shannon 提出信息论时,心理学家和其它领域的研究者们就试图将信息论的概念和方法,用于解决各自学科面临的种种难题,但因为信息科学自身的不完善以及其它学科盲目的生搬硬套,许多工作都不同程度地遭遇了困境。近年来,信息科学技术的飞速发展,以及学界对语义信息和语用信息的重新关注,使得信息科学开始肩负起联结技术科学与社会科学的重要使命。采用信息观点和技术重新阐释社会学、心理学命题的尝试,在扎扎实实地进行着。研究者们相信,对事物语义信息和语用信息的充分挖掘与利用,有助于揭示诸如思维、智慧这些高级活动的复杂过程。

心理信息学的诞生,也得益于社会科学和信息科学领域的研究者们出自不同学科视角的已有努力。最初,信息科学家们认识到只有深入理解社会科学规律,才能真正提高信息技术的社会效应,由此促成了社会计算(social computing)的诞生(Schuler, 1994; 王飞跃, 2006)。从最初的社会软件到后期的一门学科,社会计算的内涵不断演进,当下其关注的核心问题是如何结合信息科学技术和社会科学理论,对人们在互联网络世界中留下的海量真实数据进行分析,进而提高信息技术的社会应用价值(陈浩, 乐国安, 李萌, 董颖红, 2013)。

同样是利用信息技术获取、存储并分析网络世界中海量个体行为数据的计算社会科学(computational social science),则主要立足于社会科学理论视角,通过引进新的方法和技术探究新

时代背景下个体或群体间的合作与竞争关系、社会网络结构及其演化博弈等社科核心问题,属于社会科学与信息科学的交叉学科(Lazer et al., 2009)。计算社会心理学(computational social psychology)是较早就被倡导的心理学与信息技术的交叉研究领域,其主要特征是利用基于代理建模(agent-based modeling)的仿真模拟技术,探索由微观个体心理与个体间互动涌现出的宏观模式之间的非线性关系(Smith & Conrey, 2007)。

从心理学单一学科视角来看,社会计算、计算社会科学虽然带来了研究方法上的革新,但研究内容过于宽泛,计算社会心理学又难以代表网络信息时代下心理学研究的整体发展潮流,无法全面展示心理学与信息科学的密切联系和互动。作为产生和利用信息的真正智能体,人类的各种心理和行为会影响信息运动的各个环节。同时,不同信息的性质、所在载体、传递渠道和过程等,也无时不在影响着个体和群体的心理和行为。面对这样一个新交流与生活模式不断涌现、数据不断激增的崭新世界,心理学也开始面临如何处理大规模多样化数据、探索和建立新模型、建立可共享和可持续发展的数据库等方法论难题。这些都可借助信息科学所孕育的大量前沿技术在不同程度上获得解决。这预示着将信息科学技术渗透到心理学研究的各个环节和各个领域,能使双方最大程度上受益,并擦出新的科学火花。

基于此,Yarkoni (2012)在总结信息科学技术应用于心理学研究过程中的数据收集、资料整合和组织、数据挖掘和分析等方面所取得的既有成果基础之上,正式提出了“心理信息学”这一新颖的交叉学科概念。作为一个真正立足于心理学整体特色和学科发展的新学科方向,心理信息学重点关注如何在心理学研究的各个环节和各个分支领域中,借助计算机和信息科学技术促进心理学研究的与时俱进。既有研究表明,这类方法和认识上的革新不但可以帮助研究者从新视角验证心理学经典理论假设,并且在探索新的有关个体和群体的心理和行为规律方面展示出了巨大潜力,因此,它有望成为心理学未来发展的重要方向之一。

此外,需要注意的是,同为互联网时代产物的网络心理学(Cyber Psychology),虽然在研究场域上与心理信息学有所重叠,但其与几乎关照心理学整个体系和研究过程的心理信息学相较,在

研究目的、方法和内容上都有所不同。心理信息学关注的是如何应用新技术,在心理学研究整个过程发挥作用,促进心理学学科的整体发展。而目前的网络心理学则更侧重以经典理论为基础,以心理学传统范式为研究手段,强调网络是心理行为的特殊发生情境,探索与网络有关的各类心理或行为现象,属于应用心理学的分支方向之一。

## 2 研究过程中的方法与工具革新

心理学自诞生伊始就以科学认识论与方法论为学科指导原则,一直在研究方法上进行探索和改进,努力形成严谨、成熟的学科研究范式。但与此同时,心理学研究的科学性不断受到来自学科内外的诸多质疑,究其原因主要来自于研究方法和技术上的局限。其中,取样缺乏代表性和研究手段人为性问题比较突出:受人力、时间、经济成本和可操作性限制,传统心理学研究多以大学生为被试,即使是大规模调查,也很难真正意义上代表总体。另一方面,心理学的核心目标是探索现实生活中人类心理和行为的因果机制,但典型的实验法研究通常在毫无“人情”的实验室控制情境中进行,高度抽象并明显脱离实际生活,这无疑降低了心理学研究的外部生态效度,制约着心理学的进一步发展(Gosling, Gaddis, & Vazire, 2008; 彭凯平, 刘钰, 曹春梅, 张伟, 2011)。

网络时代的到来以及计算机和信息科学技术的飞速发展,给心理学研究带来了福音,以研究方法和技术工具的深刻变革为重要特征的心理信息学,为心理学研究过程的各个阶段注入了新活力。

### 2.1 数据收集

无论是以互联网为平台进行调查和实验,还是利用移动设备收集各类心理行为数据,甚至是直接抓取网络生态数据信息,都使心理学研究数据在数量和质量上有了很大提升。

#### 2.1.1 网络实验和调查

相对于其它研究方法,网络实验和调查对心理学研究者来说并不陌生,特别是在西方国家,已经出现了很多方便学者们进行研究的专门化网站。譬如,“调查猴子”(SurveyMonkey)和 Qualtrics 就设计了一些易于操作的应用程序,以便人们能够轻松地创建和管理复杂的研究。为充分发挥网络实验的模拟性和便利性,这些网站还允许研究者利用简单的 JavaScript 等编程语言自行设计和

实施与现实情境相似的实验。此外,“亚马逊土耳其机器人”(Amazon's Mechanical Turk, MTurk)等网站的出现,则使得招募高质量被试变得更加容易和经济。该类网站包含了完整的参与者报酬系统、大规模被试库,以及从研究设计、招募被试到数据收集的整个简易流程。

面对同行关于网络实验、调查信效度如何的忧虑和疑问,一些研究者针对利用 Mturk 等网站招募的样本人群特征和收集的数据质量进行了分析。结果发现,与传统研究样本相比,网络样本在性别、年龄、地区和社会经济地位等方面都更具区分性和代表性;虽然人们决定是否参与研究受报酬水平和任务长度的影响,但通过这些网站仍能快速而经济地招募到被试;实际的报酬水平并不会影响到数据的质量;收集到的数据的可信度至少不会低于传统方法,而且网络调查结果的外部效度不受重复作答者或不认真作答者的影响(Gosling, Vazire, Srivastava, & John, 2004; Buhrmester, Kwang, & Gosling, 2011)。互联网作为一种新的研究工具,其价值不仅体现在扩展了心理学的研究平台和实验范围,更重要的是,其为解决传统研究的弊端提供了一个可能的途径。

#### 2.1.2 移动设备的应用

除了互联网以外,研究者们对移动设备(如手机)的关注和使用也在不断增多。很多年前,他们就开始使用各类移动设备记录和收集数据。有研究者通过电信服务商提供的用户电话拨打时间、长度、号码和所处位置等信息,追踪人们的社会联系,进行社会网络分析(Campbell et al., 2008; Eagle, Pentland, & Lazer, 2009)。限于数据的匿名性,那时研究者还未能充分利用到移动设备的传感性、联结性和互动性等功能。

另外一些研究者通过购买或设计特定功能的电子记录设备,收集相对少量被试较为详细的心理和行为数据。譬如,一种叫做“电子激活录音器”(Electronically Activated Recorder, EAR)的便携式记录设备(现已开发为 iPhone 手机的 APP),能定期记录日常生活中人们的声音信息,帮助研究者了解到真实情境中的人类行为(Mehl, Pennebaker, Crow, Dabbs, & Price, 2001)。性别差异一直是心理学关注的议题。有关女性更加多语健谈的刻板印象,已深深根植于东西方各类民谚和流行文化中,且被认为是不证自明的科学事实。其实之前

很长一段时间,并没有任何研究系统地记录过大批人群的自然交谈过程,为这一假设提供直接可靠的数据检验。有研究者曾通过分析153个英国公民日常对话的磁带录音推断,女性平均每天说8805个单词,男性平均每天说6073个单词。但该研究者也承认,这种估计存在问题,因为他缺少有关被试何时关闭录音机的信息(Liberman, 2006)。为更加科学地检验这一假设,研究者通过使用EAR来对人们在现实生活中的即时互动,进行不加干扰和不可自控的记录。由此得到的结果是,男性和女性平均每天说16000个单词,性别间差异不显著,但个体间差异显著(Mehl, Vazire, Ramirez-Esparza, Slater, & Pennebaker, 2007)。Vazire和Mehl(2008)做的另一项研究发现,与自己关系密切的人往往比个体本人更能预测该个体的日常行为。这就提醒广大研究者,通过自我报告评估人们的心理和行为,甚或是人格特质,也许远非想象中那般可靠。虽然基于新设备的研究发现了许多令人深省的研究结果,但这些设备大多昂贵、笨重,且需要实地分发给被试,因此在学界并未获得推广。

近年来,随着智能手机的大幅普及以及受到手机应用程序开发者的启发,一些心理学研究者着手开发能够供网络下载的手机应用软件,这类软件通常能够自动管理知情同意、资料收集、资料上传和费用支付等研究过程,已在心理学研究中崭露头角(Miller, 2012)。Dufau等人(2011)指出,传统认知心理学研究利用小规模同质自愿者群体,考察注意、记忆、语言等人类认知功能所存在的局限性和可能的取样偏差,进而展示了将手机作为收集大规模异质人群数据的工具的可能性和优势,以及对整合认知领域相关研究与理论的巨大潜力。其和同事开发的一种多语言版本语言任务决策应用程序,在4个月的时间内收集到4000名参与者的数据。Eagle等人(2009)为解决以自我报告为主要数据源的人际互动和社会网络结构研究,在深度、广度和准确性等方面所面临的瓶颈,发起了名为“现实挖掘”(reality mining)的研究项目。该项目通过分析94位智能手机用户连续9个月的位置和通话记录,以及有关日常行为和人际关系的自我报告,发现两种数据源的分析结果虽有重叠,但并不完全一致,口头报告更容易随互动时间和显著性的变化而发生偏移。并且,结果显示

仅仅依靠手机客观数据就能很准确地推测出人们之间的朋友关系,甚至能够预测诸如地位、工作满意度等个体水平层面的社会与心理变量。Killingsworth和Gilbert(2010)通过一个基于iPhone手机的应用软件,创建了一个大范围人群日常生活中的即时思维、感受和行动报告的大规模数据库。在此基础上,研究者发现人们经常处在“走神”(wondering)状态,且不管此时正在干什么,都会伴随不愉快的情绪体验。人们在想什么比人们在做什么更能预测人们的情绪状态。该研究验证了走神实际上是一种认知进化但却以情绪消耗为代价的观点。

### 2.1.3 网络生态数据的抓取

互联网的出现深刻地改变了人们的生活方式:虚拟和现实的界限愈加模糊,人们通过网络交流信息和情感,更新自己的近况和活动,上传和分享图片,下载音乐和电影,购买商品和寻求问题答案。计算机和信息科学技术的发展,使得这些反映现代生活印迹的海量数据得以永久保存,并可通过一定程序或渠道获得,为心理学家探究人们的心理生活提供了无以比拟的大规模生态数据库。

从起初通过电子邮件分析探索社交网络特征(Dodds, Muhamad, & Watts, 2003; Adamic & Adar, 2005),到随后基于博客的情感分析技术来探索人们的情绪变化趋势(Balog & Rijke, 2006),及其与现实行为的相关性(Mihalcea & Liu, 2006; Mishne & Glance, 2006),再到近年来针对Facebook和Twitter等著名社交媒体网站上海量用户数据的集中挖掘(Dodds, Harris, Kloumann, Bliss, & Danforth, 2011; Gosling, Augustine, Vazire, Holtzman, & Gaddis, 2011; Lansdall-Welfare et al., 2012; Zhang, Fuehres, & Gloor, 2011)。研究者对网络媒介的关注和把握正是为了映射人们的真实生活,网络发展和技术进步使这一看似单纯的目标得以实现。

### 2.2 数据的组织与整合

为显著提高数据收集和分析工作的学术共同体利益,心理学研究者们正在积极构建许多分支研究领域下可共享和重复利用的数据库和数据分类体系,以便对分支研究领域乃至整个学科的多源头数据进行迭代重组与整合工作。

#### 2.2.1 数据库的建立

以研究工具技术创新著称的认知神经科学,可以说是实践心理信息学主旨的典范,该领域的

研究者已经不再局限于个别研究的新技术吸纳与创新,而是立足于整个学科领域的发展。一些研究者和研究机构开始致力于开发大规模在线数据库,以促进跨领域数据的整合和分享。譬如最近建立的一个公共数据库“千人功能连接组脑项目”(1000 Functional Connectomes Project),包含了来自35个地区的1400多名被试的大脑功能扫描数据,为研究者提供了非常有价值的大规模人类静息态数据(Biswal et al, 2010),有利于对人脑功能的协同深入探索。脑地图数据库(BrainMap Database)和界面管理系统数据库(Surface Management Systems DataBase, SumsDB)等还提供了快速数据挖掘技术、可视化技术、整合与分析不同研究中立体坐标技术(如检索和神经影像分析等),从而为研究者在不同研究领域间建立连接,验证已有假设,以及开发新领域、发展新假设和理论等方面提供帮助(Laird, Lancaster, & Fox, 2005; Dickson, Drury, & Van Essen, 2001; Yarkoni, Poldrack, Van Essen, & Wager, 2010)。

数据库应用的另一重要领域是在线文本情感分析。它通常建立在特定情感词库或语料库基础之上,通过词汇匹配技术(term-based matching technique),将文本中的词汇与情感词库或相应语料库中的词语进行匹配,从而判断文本在情感方面的倾向或极性(O'Connor, Balasubramanyan, Routledge, & Smith, 2010)。研究者们已经通过人工、自动或半自动化等方式建立了一些情感词库,如标准英语情感词库(Affective Norms for English Words, ANEW)、意见发觉者(OpinionFinder)、情感词语网络词典(WordNet-Affect)和谷歌心境状态量表(Google-Profile of Mood States, GPOMS)等(Bradley & Lang, 1999; Wilson, Wiebe, & Hoffman, 2005; Strapparava & Valitutti, 2004; Bollen, Mao, & Zeng, 2011)。

此外,很多类似于情感词库的特定领域数据库,以及其它一些有关人们心理与行为的数据库,都可在特定社会科学网站,如哈佛定量社会科学研究所(Institute for Quantitative Social Science, IQSS)的The Dataverse Network或专业领域数据库,如开放性功能核磁共振项目(OpenfMRI)中下载。这些共享资源有助于研究者对彼此的研究结果进行比较、验证、扩展和整合,加速该领域的研究进程。

## 2.2.2 分类体系的建立

有心理学研究者担忧如果任凭各类数据库各自为战地独立建设,可能会导致杂乱无序的学术研究状态,反而会阻碍各个领域乃至整个学科的发展。因此,为最大限度利用公共数据库并更有效地对数据进行结构化访问与搜索,一些心理学家尝试建立全面的心理结构分类体系。Poldrack等人(2011)的“认知地图集项目”(Cognitive Atlas)就试图建立一个能够被学界广泛认可的认知心理学概念与研究任务的智识体系,从而促进相关研究的自觉整合。Balota, Yap, Hutchison 和 Cortese (2013)进一步利用大规模数据挖掘技术,对三个独立数据库(视觉词语再认数据库、语言启动数据库和再认记忆数据库)进行了统合分析,尝试建立起能够被共享和公认的有关语言加工过程和关键影响因素的普适原则。这些工作充分体现了大规模数据研究在建立共享数据库、探索各变量间有机、有序联系,以及建立连续性理论体系中的重要性和必要性。

相关工作刚刚开始,距离建立起一个容纳所有心理学分支的完整分类体系还有很长的路要走,但是相应数据库与基础分类体系的完备,的确会极大促进心理学研究者组织、分享和获取学术共同体协同努力的成果,从而加速对人类心理世界的探索之旅。

## 2.3 数据分析

计算机和信息科学技术的发展,给心理学带来的不仅仅是对海量社会生态数据的捕获,更是一场分析工具和技术上的变革。一方面,各类开源统计建模软件(如R软件)和Ucinet、KrackPlot、Negopy等网络分析软件,使很多曾经不能实现的分析变成了可能,推动了相关研究进程。另一方面,作为心理信息学数据分析的核心技术,计算机数据挖掘技术的逐步渗入,无形中促进了心理学探索性研究的发展,补充了传统心理学以假设验证为主要特征的研究逻辑,有助于产生全新的假设和理论。

一些研究者开始利用各种计算机算法进行数据挖掘,揭示大规模数据视野中人类心理与行为的新特征和潜在新关联。譬如,Yarkoni和他的同事开发了一个叫做“神经合成”(Neurosynth)的计算平台,该平台利用数以千计的神经影像论文中的数据,自动生成fMRI元分析结果,并且能对大

脑活动模式进行自动解码(Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011)。计算机技术在数据分析中的另一重要应用,是利用机器学习算法对研究数据进行分类和预测。Pereira, Mitchell 和 Botvinick (2009)曾引入机器学习算法对大脑活动状态和心理失调症状进行编码、分类和预测。这些算法有助于解决长久困扰学界的过度适应问题(Kriegeskorte, Simmons, Bellgowan, & Baker, 2009)。

对人类社交网络的拓扑结构分析,更是以各类计算机算法为基础,如马尔科夫法(如 PageRank 算法、HITS 模型)、度中心法(degree centrality)、基于路径法(如  $\alpha$ -centrality、SenderRank)等(孟小峰, 余力, 2011)。这些数据挖掘技术的应用使研究者能够更好地探究人与人之间的彼此关系和相互影响,有助于解决复杂网络社会的非线性涌现问题(程学旗, 陈海强, 韩战钢, 2006)。

正如 Yarkoni (2012)所言,“如同基因组关联研究给复杂的疾病基因学带来的革命那样,这种大规模数据库、丰富的样本和经验数据与计算机算法的结合,可能会使我们很快识别出各种心理和行为现象中的新相关……发现一些传统研究方法无法发现的反直觉性结果”。

## 2.4 结果的评估与分享

心理信息学不同于心理学其它交叉或分支领域的一个显著特征,体现在其对心理学学科健康发展的整体关注上。研究者们重视利用信息技术促进心理学研究的各个环节都能取得实质性进步,其中也包括如何提升研究成果的发布质量,以及如何促进研究成果在领域内外的共享与交流。

为此,该领域的先驱们提出了一些供学界同仁参考的有益方向:通过引入商业和社交网站中广泛使用的协同过滤算法,提高对已发表文章的评论质量;通过自动化质量控制算法,检测学术出版物中可能存在的偏见;开发专门用于检索未发表论文的数据库;设计个性化推荐技术,帮助个体研究者识别他可能感兴趣的文章;利用已有媒介或开发新的专门化媒介,促进研究成果在领域内的讨论以及面向大众的分享和普及(Yarkoni, 2012)。

## 3 传统研究内容的新拓展

相较于遵循类同理路出现的经济信息学、生物信息学和神经信息学等交叉学科的快速发展和相对成熟,心理信息学还刚刚起步,以它为明确

主题的研究还比较少。但不可否认的是,心理学和信息科学在思想和技术上的融合,已经在情绪、人格等个体心理领域,以及社会网络和群体动力等社会(群体)心理领域给我们带来了不少惊喜。

### 3.1 个体心理领域

计算机和信息科学技术的应用,使研究者可以从微观和宏观层次上共同探索个体心理规律,揭示其对个体行为和社会生活等方面带来的影响,其中比较典型的研究领域是情绪和人格。

#### 3.1.1 情绪研究

不管是单纯为了基础科学研究,还是为了社会或商业应用,情绪一直是心理学和信息科学研究的重点之一。得益于数据收集方法及文本情感分析技术的发展,研究者在此领域已取得了一些令人瞩目的成果,在微观或宏观层面上为我们描绘出人类情绪的新图景。

##### (1)情绪的内在机制和外表达

在微观层面上,研究者诉诸于大规模真实数据和挖掘技术,探索情绪的内在机制和外表达。譬如,以基于内核图像重建为核心的标准化元分析技术,使得研究者可以对数十、甚至数百个研究中的神经影像结果进行整合性分析,已被应用于情绪失调和脑功能特征的关系研究中。Wager 等人(2008)通过该元分析技术发现,杏仁核不仅与害怕情绪有关,在厌恶情境下也会表现出不同程度的激活;杏仁核和脑岛在三种不同焦虑状态下都有较高水平的激活,但是创伤后应激反应只引起腹内侧前额叶皮层的高度激活。这些发现充实着我们对情绪内在机制的理解(Etkin & Wager, 2007)。Kober 等人(2008)指出,以往负性情绪模型只关注杏仁核的作用,忽视了动物模型研究中相关脑区发现,如中央灰质和丘脑下部。元分析显示,这两个区域在消极情绪时会出现重复性激活现象,因此有必要对两者的功能和重要性进行重新考量。

另一方面,为更有效帮助情绪障碍患者、协助心理卫生工作者,一些研究者探索利用社交媒体诊断或预测个体的抑郁障碍发作,并获得了意外收获。例如,有研究者采用经济省时的网络数据抓取技术和 MTurk 上获得的抑郁量表测试数据构建样本库。通过分析这些个体在 Twitter 上的发帖表达特征,研究者发现患有抑郁障碍的被试,在确诊前一年的社交活动就开始减少,消极情绪

体验有所增加,同时表现出更多的自我关注和对抗抑郁药物的关注等。基于这些线索,研究者们建立了一个统计分类指标,预测人们罹患抑郁障碍的可能性。实践证明,该指标的预测准确率达到70%以上(De Choudhury, Gamon, Counts, & Horvitz, 2013)。

### (2)大众情绪变化规律和情绪指标

海量数据分析技术也让研究者可以从微观至宏观的路径,把握由若干个体叠加起来的大众情绪整体特征、变化规律,及其与现实社会经济事件的相关性,从而建立具有预测效力的宏观情绪指标体系。

Golder 和 Macy (2011)曾对数百万的 Twitter 使用者的在线信息进行分析,将他们的情绪变化趋势以天或周为单位进行量化,在个体季节性情感障碍成因的两个竞争性假设中进行甄别验证。Dodds 和 Danforth (2010)采用人们对大量词汇的情绪效价评估结果,对不同类型的大规模文本,如歌曲名和歌词、网络博客、国情演说等,进行数据挖掘。结果发现,从20世纪60年代到20世纪90年代中期的歌词中所传达的幸福感水平有逐渐下降的趋势,博客中所表达的幸福感水平在2005年至2009年间稳步上升。为更全面地了解公众在 Twitter 上的情绪表达规律,研究者扩大了词汇的选取范围,增加了词库容量,并采用与 ANEW 相类似的方法,建构了一个可调节的、实时的、免干扰的和基于文本的测量工具。该方法的核心是人们对一系列词汇的幸福度评估以及一个将词汇范畴扩展至所涉猎文本的新算法。在此基础上,Dodds 等人(2011)通过分析6300多万用户在33个月间所发的约46亿条 Twitter 信息,揭示出了大众幸福感水平在不同时间尺度上(几小时到几年)的变化规律。

为探索情绪和现实事件之间的关联性,有研究者对1084名美国公民在911前后各两个月间的在线日记文本进行了分析,结果显示,那些有更多认知性和社会性卷入的参与者表现出了更强烈的消极情绪,但这些情绪在事件发生两周后都恢复到正常水平(Cohn, Mehl, & Pennebaker, 2004)。Back, Küfner 和 Egloff (2010)对911事件当天的约50万条手机短信进行了分析,揭示出了不同类型情绪的变化规律:人们最开始感到愤怒,随后变得焦虑,但很快恢复到正常水平,最后出现的是

悲伤情绪。

研究者还尝试建立能够预测现实世界中的群体行为或社会经济事件的情绪指标。如 Devitt 和 Ahmad (2007)曾通过识别金融评论文本的情感极性,预测金融市场的未来走势。Kim 和 Hovy (2007)通过分析大样本网络新闻评论预测美国大选结果。Bollen, Pepe 和 Mao (2011)利用 OpinionFinder 和 GPOMS,分析人们在 Twitter 上的特定类型情绪表达,发现了它们是能显著提高预测道琼斯工业指数(Dow Jones Industrial Average, DJIA)效果的情绪指标。Mao, Counts 和 Bollen (2011)在 Google 和 Twitter 上搜索与股市相关的信息,将之与相应的情绪指标进行比较,发现“谷歌搜索观察”(Google Insight for Search, GIS)和“Twitter 投资者每日情绪指数”能预测股市波动。Lansdall-Welfare 等人(2012)利用 WordNet-Affect 分析了98万英国人在 Twitter 上历时31个月所发布的信息,发现大众的某些情绪能预测社会事件的发生,如在2011年夏季的伦敦暴动发生之前,公众的愤怒情绪从春季就开始持续增加。这些研究昭示了信息技术与情绪理论相结合在应用上的巨大潜力。

### 3.1.2 人格研究

人格作为反映个体稳定心理行为倾向的重要特征,一直受到心理学和诸多领域的关注。目前,研究者们开始考虑如何利用、开发新的方法和技术,促进人格特质的测量及其相关影响因素研究的进一步发展。

Yarkoni 曾开发一种能够以最小代价自动化缩短各种人格问卷的方法,并实证验证了该方法的有效性。在此基础上,其生成了一个包含181道题目的新量表,该量表能够准确地获得原本由203个不同量表组成的8个新的不同问卷的得分(Yarkoni, 2010a)。有研究者利用 Twitter 的 API 收集使用者的个人资料,然后结合已有人格测量结果,采用两种机器学习算法进行训练,以期通过用户个人资料信息预测他们在大五人格维度上的得分(Golbeck, Robles, Edmondson, & Turner, 2011)。

Quercia 和其团队分析了不同类型 Twitter 用户的人格特征。他们发现不同类型用户并非所谓相似“节点”,而是既有共性又具个性的独立个体。例如,受欢迎的用户和有影响力的用户,虽然都表现出外向性和情绪稳定性的共同特征。但是受欢迎的用户在开放性上的得分普遍更高,更富于

幻想;而有影响力的用户则表现得更具组织性,其消极的情绪表达也更为明显(Quercia, Ellis, Capra, & Crowcroft, 2011; Quercia, Kosinski, Stillwell, & Crowcroft, 2011)。Gosling 等人(2011)通过分析 Facebook 上的数据,探索人们在社交网站中表现出的人格特征、自我报告网络行为和客观描述性信息之间的相关。结果发现,外向性不仅能够预测人们使用社交网站的频率,而且能够预测其卷入程度和活动强度。对朋友列表和图片帖子的分析表明,外向性的人更倾向于在现实生活和虚拟生活之间建立联系,也就是说,不同于大众的猜测,在线社交网站使用者不是通过虚拟网络行为逃避或补充其现实生活中的人格,而是将其线下的人格特征拓展至网络世界中。

Yarkoni (2010b)指出,以往人格特质和个体网络用词特征间的相关研究,受方法和技术局限,只探讨过主要人格类型和整体词语类别间的相关,对内在具体关联甚少涉猎。为此,他搜集了 694 名 Google 博主的近 80 万字文本数据,系统分析了人格、整体词语类别和个体使用词汇特征间的相关,所得结果既验证了以往现实情境中的相关研究结论,又发现了许多新的涉及分类水平、单个词语使用与人格特征的相关关系。

这些研究不但有助于验证心理学相关理论、探索新发现,并且对推荐系统、市场营销定位和网站界面设计等实际应用也有重要启发。

### 3.2 社会心理领域

不同于个体水平,社会水平的心理研究侧重于分析特定社会情境或社会结构条件下的心理现象,强调的是社会互动的作用和影响。群体心理和行为并非群体内所有个体心理和行为的简单相加,而是群体成员在社会互动情境中所形成的共有或典型的价值观、态度和行为方式。社会心理的广泛影响以及新的网络社交环境与形态的出现,吸引了众多研究者涉足该领域。社会心理研究对样本代表性和生态效度的长期高诉求,也使得该领域对新技术的应用具有天然敏感性。短短几年里,相关研究已涉及虚拟和现实网络社会中的诸多现象。

#### 3.2.1 虚拟网络与社会心理

随着 Web 2.0 理念的深入人心,人类社会迎来了一个崭新的网络社交时代。据 2010 年 2 月数据显示,全球最大的社交网站 Facebook 的注册用

户已达 4 亿多,成为排在中国和印度之后的人口世界第三大社会(King, 2010)。虚拟网络社会不但聚集了大量用户,而且也改变了人与人之间的互动模式。人们通过各种网络工具与他人沟通交流,结成不同类型的关系,如社交网络中的好友关系、微博中的关注关系、购物网站或论坛中的共同兴趣关系等。以往研究表明,网络中的社会关联普遍存在局部聚集特性(程学旗,沈华伟, 2011)。大量具有局部聚集特性的关系数据,启发着研究者们借助计算机和信息技术,揭示特定虚拟网络环境中所承载的社会心理与行为规律。

Szell, Lambiotte 和 Thurner (2010)基于数据挖掘和社会网络分析技术,对一个容纳 30 多万玩家的大型在线游戏中的多维社会网络进行了整体性结构分析,并根据个体间不同的交互关系,抽取了 6 种不同的社会网络,探索了不同类型的网络结构特性以及它们之间的关联和重叠关系。结果发现,个体在不同的网络中倾向于扮演不同的角色,并验证了用以解释社会冲突和紧张关系的结构平衡理论。

另有研究者尝试通过分析 Twitter 用户的人际地理拓扑结构,挖掘影响人们使用社交网站的动机。结果显示,人们使用 Twitter 主要是为了日常交流、搜寻和共享信息,而具有相同需求的人更容易通过彼此联系而结成社群(Java, Song, Finin, & Tseng, 2007)。从外部动机上看,驱动人们使用社交网络的并不是那些外显的朋友和同事关系,而是潜藏其下的、零散的网络社交关系(Huberman, Romero, & Wu, 2008)。另外,研究者还发现用户发帖特征和网络历史使用率,与用户背景信息相比,能更好地预测该用户所发信息的网络扩散程度(Yang & Counts, 2010)。

这些现象虽然多发生于虚拟网络世界中,但对于检验已有的基于现实世界现象发展而来的社会心理理论的普适性,以及对于理解现实和网络交织的大时代背景下,人们的社会生活新特征和心理行为规律等,都具有重要意义。

#### 3.2.2 社会网络结构与社会心理

除对虚拟情境中社交生活的关注以外,研究者们同时尝试揭示现实社会网络结构对个体和群体行为的影响。Kossinets 和 Watts (2006)通过记录一个学年的电子邮件标题信息,考察了某大学 43,533 个学生、教师和职工组成的动态社会网络。



结果发现,该社会网络的发展主要受网络形态自身以及其所处组织结构的影响。当整体扰动消失后,网络特征的平均水平接近平衡,而个体特征则并不稳定,该发现有助于推测网络中个体间的交互作用。

为解答“幸福感能否在人际网络中扩散?”以及“幸福感是否会在人际网络中集群?”的问题,Fowler和Christakis(2008)开展了一项纵向研究,他们收集了4739个被试20年间的相关数据,涉及不同类型社会网络和社会联结。该研究结果显示,现实社会网络中确实存在明显的幸福集群和不幸福集群,并且被许多幸福的人包围着的个体,或是处于幸福集群中心的个体更可能变得幸福。纵向统计模型表明,幸福集群的形成主要是幸福感传播的结果,而不是由于相似幸福感水平的人更倾向于聚集在一起所致。有研究者对2008年9月到2009年9月间发布在Twitter上的近4千万条消息的分析发现,个体的平均幸福感体验水平与其在网络中的一个、两个和三个联结点外的用户幸福感得分皆有显著的正相关,从而再次证明了网络形态对幸福集群的影响(Bliss, Kloumann, Harris, Danforth, & Dodds, 2012)。

在合作行为问题上,人们通常认为合作行为会随着个体间互动的增加而增加。依社会网络视角而言,这意味着高聚集的社会网络中的合作行为要多于低聚集的社会网络。为检验该假设,Suri和Watts(2011)开展了一系列基于真实社会网络的实验研究,结果发现网络结构形态对整体合作行为水平的平均贡献率没有显著影响。合作行为只在某些情况下对网络中的直接邻居有正向感染力。另有研究发现,动态网络中的合作行为存在着局部强化和全局扩散间的权衡关系:当网络中建立联系的代价大,并且局部结构大规模消失时,更容易产生高水平的合作行为;当网络中建立联系的代价小,并且朋友的朋友之间的互动可能性高时,建立高水平合作行为的可能性反而更小。对这一反直觉性结果的可能解释是,人们能够单方面地切断与他人的联系,而新的联系只有在双方都同意的情形下才能建立(Hanaki, Peterhansl, Dodds, & Watts, 2007)。

#### 4 问题与展望

综上所述,目前计算机和信息科学技术的应

用已经对心理学研究的设计思想、数据资料的收集、组织和整合、结果的挖掘分析等几乎各个环节带来了发展,已渗透到心理学的诸多分支研究领域。在研究方法上,心理信息学利用各种平台和技术收集各类心理行为信息的能力,以及对大规模多元真实数据的分析处理能力,显著提高了研究效率和研究结果的生态效度,极大地扩展了研究视角和研究范畴。在研究内容上,虽然心理信息学立足于心理学科的全面发展,但目前的研究多集中在探讨基于海量数据的心理和行为变化规律,多变量间相关性分析,以及社会心理学所关注的人际互动和社会网络结构对个体、群体心理的影响。随着心理信息学的研究范畴向认知神经甚至政治、经济等多层次相关现象领域的拓展,其验证和发展心理学理论假设的潜力将获得更大程度的发挥。

但是,作为一门新兴的交叉学科,学界对于心理信息学的认识和应用远未成熟和普及,还有许多值得注意和亟需解决的问题。

从技术层面上看,目前对大规模数据的获取、存储、整理和分析的具体技术还有继续发展和完善的空间。例如,如何避免其它因素的干扰,提高手机等移动设备收集人们日常心理和行为数据的质量和效度(Miller, 2012);如何改进情感分析技术,使之能够对不同语境下的词语属性,对整体文本所涉的评价对象的具体态度极性等进行更为智能的识别判断(赵妍妍,秦兵,刘挺,2010);如何整合分析网络中的文本、图片和声音等异质数据。但是,这些技术问题的解决并不只是某一学科的份内任务,而是在信息科学技术和多学科理论视角的不断撞击中才能实现。

从研究范式上看,虽然新技术的应用对于改善传统心理学研究中的取样缺乏代表性和研究手段的人为性等弊端提供了可能性。但目前,该领域的许多研究仍存在着研究思路较为简单,研究设计缺乏严谨性等缺陷。少有研究通过严格的实验控制设计进行因果推断,目前的绝大部分研究只限于相关关系探讨。因此,有必要在未来的研究中着重结合心理学业已成熟的实验范式,充分利用新技术和大规模生态数据揭示人类心理和行为的本质规律。

从研究伦理上看,新的方法和技术创新带给人们的不仅仅是兴奋,还有对信息安全和个人隐

私等问题的重新思考和担忧。一方面,很大一部分涉及个人隐私(特别是消费者隐私)的数据为某些私营公司或个别机构所有,其对数据安全性的理解,以及有关数据共享的协议并不统一。这可能在使用和分享私人信息时,造成对个人或某些机构的威胁和伤害。另一方面,许多原本看似匿名的信息,如果综合多渠道来源数据或利用特殊技术分析,仍然会泄露很多个人的真实信息。因此,有必要建立一套自上而下的监管机制,使得从数据信息的集中收集、存储,到信息在商业和学术界的共享和使用,都有健全有效的制度可依,从而保证个人信息安全,并促进学术和应用研究的健康发展(Cioffi Revilla, 2010)。

从研究重心和人才培养上看,一方面,为了防止心理学陷入唯技术工具驱动的泥潭,以至将心理学埋葬在纷繁复杂的数据之中(Karpf, 2012)。心理学研究者,特别是有志从事心理信息学研究的学者,应本着以问题为中心的原则,重视新技术和传统研究方法的结合,在确定有意义的研究目标和内容基础之上,充分利用新的数据收集和分析技术,如从封闭化的软件(如 SPSS 软件)到改为使用开源软件(如 R 软件),以及花费一定时间和精力研习机器学习等数据挖掘技术。对于已具备相关信息科学技术背景的心理学者,则可以尝试寻找能够与信息技术相结合的基础或商业研究项目。另一方面,由于现有技术的高端性和复杂性,使得传统心理学研究者一时难以介入其间。以往的许多心理信息学研究都是计算机和信息科学领域专家开展的,其专长于应用研究和具体算法的改进与创新,对心理和行为理论的系统研究所涉不深。因此,为使心理信息学乃至整个心理学获得长足发展,需要在心理学学科人才的培养计划中,增加对在校学生的相应学术训练,重视培养同时擅长心理学理论分析和信息科学技术的综合性人才,并鼓励跨领域研究成果的发表与分享。

## 参考文献

- 陈浩, 乐国安, 李萌, 董颖红. (2013). 计算社会科学: 社会科学与信息科学的共同机遇. *西南大学学报(社会科学版)*, 39(3), 87-93.
- 陈惟昌, 王自强, 陈志华, 安荣妹. (2001). 神经信息学的原理与展望. *生物物理学报*, 17(4), 613-620.
- 程学旗, 陈海强, 韩战钢. (2006). 社会信息的网络化分析初探. *中国计算机学会通讯*, 2(2), 18-26.
- 程学旗, 沈华伟. (2011). 社会信息网络中的社区分析. *中国计算机学会通讯*, 7(12), 12-20.
- 孟小峰, 余力. (2011). 用社会化方法计算社会. *中国计算机学会通讯*, 7(12), 25-30.
- 彭凯平, 刘钰, 曹春梅, 张伟. (2011). 虚拟社会心理学: 现实, 探索及意义. *心理科学进展*, 19(7), 933-943.
- 邱均平. (2002). *市场经济信息学*. 武汉: 武汉大学出版社.
- 王飞跃. (2006). 社会计算的意义及其展望. *中国计算机学会通讯*, 2(2), 28-35.
- 张阳德. (2009). *生物信息学(第二版)*. 北京: 科学出版社.
- 赵妍妍, 秦兵, 刘挺. (2010). 文本情感分析. *软件学报*, 21(8), 1834-1848.
- 钟义信. (2002). *信息科学原理(第三版)*. 北京: 北京邮电大学出版社.
- Adamic, L., & Adar, E. (2005). How to search a social network. *Social Networks*, 27(3), 187-203.
- Back, M. D., Küfner, A. C., & Egloff, B. (2010). The emotional timeline of September 11, 2001. *Psychological Science*, 21(10), 1417-1419.
- Balog, K., & de Rijke, M. (2006, May). *Decomposing bloggers' moods: Towards a time series analysis of moods in the blogosphere*. Paper presented at the WWW-2006 Workshop on the Weblogging Ecosystem, Edinburgh, Scotland.
- Balota, D. A., Yap, M. J., Hutchison, K. A., & Cortese, M. J. (2013). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Ed.), *Visual Word Recognition Volume 1: Models and methods, orthography and phonology* (pp. 90-115). New York, NY: Psychology Press.
- Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., ...Milham, M. P. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10), 4734-4739.
- Bliss, C. A., Kloumann, I. M., Harris, K. D., Danforth, C. M., & Dodds, P. S. (2012). Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *Journal of Computational Science*, 3, 388-397.
- Bollen, J., Mao, H. N., & Zeng, X. J. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2, 1-8.
- Bollen, J., Pepe, A., & Mao, H. (2011). *Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena*. Paper presented at the Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, Barcelona, Spain.
- Bradley, M. M., & Lang, P. J. (1999). *Affective norms for*

- English words (ANEW): Instruction manual and affective ratings (Technical report C-1)*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.
- Brillouin, L. (1956). *Science and information theory*. New York: Academic Press.
- Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1), 3–5.
- Campbell, A. T., Eisenman, S. B., Lane, N. D., Miluzzo, E., Peterson, R. A., Lu, H., ... Gahng-Seop, A. (2008). The rise of people-centric sensing. *Internet Computing, IEEE*, 12(4), 12–21.
- Cioffi-Revilla, C. (2010). Computational social science. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 259–271.
- Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science*, 15(10), 687–693.
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). *Predicting depression via social media*. Paper presented at Proceedings of the 7th International AAAI Conference on Weblogs and Social Media, Boston, MA.
- Devitt, A., & Ahmad, K. (2007). *Sentiment polarity identification in financial news: A cohesion-based approach*. Paper presented at the Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, page 984–991, Prague, Czech Republic.
- Dickson, J., Drury, H., & Van Essen, D. C. (2001). 'The surface management system'(SuMS) database: A surface-based database to aid cortical surface reconstruction, visualization and analysis. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 356(1412), 1277–1292.
- Dodds, P. S., & Danforth, C. M. (2010). Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. *Journal of Happiness Studies*, 11(4), 441–456.
- Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS One*, 6(12), e26752.
- Dodds, P. S., Muhamad, R., & Watts, D. J. (2003). An experimental study of search in global social networks. *Science*, 301(5634), 827–829.
- Dufau, S., Duñabeitia, J. A., Moret-Tatay, C., McGonigal, A., Peeters, D., Alario, F.-X., ... Grainger, J. (2011). Smart phone, smart science: How the use of smartphones can revolutionize research in cognitive science. *PLoS One*, 6(9), e24974.
- Eagle, N., Pentland, A. S., & Lazer, D. (2009). Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, 106(36), 15274–15278.
- Etkin, A., & Wager, T. D. (2007). Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *The American Journal of Psychiatry*, 164(10), 1476–1488.
- Fowler, J. H., & Christakis, N. A. (2008). The dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study. *British Medical Journal*, 337, a2338, 1–9.
- Golbeck, J., Robles, C., Edmondson, M., & Turner, K. (2011). *Predicting personality from twitter*. Paper presented at the Privacy, Security, Risk and Trust (Passat), 2011 IEEE third international conference on and 2011 IEEE third international conference on social computing (SocialCom), 1–23.
- Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science*, 333(6051), 1878–1881.
- Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking*, 14(9), 483–488.
- Gosling, S. D., Gaddis, S., & Vazire, S. (2008). First impressions based on the environments we create and inhabit. In N. Ambady & J. J. Skowronski (Eds.), *First Impressions* (pp. 334–356). New York: Guilford.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist*, 59(2), 93–104.
- Hanaki, N., Peterhansl, A., Dodds, P. S., & Watts, D. J. (2007). Cooperation in evolving social networks. *Management Science*, 53(7), 1036–1050.
- Huberman, B. A., Romero, D. M., & Wu, F. (2008). Social networks that matter: Twitter under the microscope. *First Monday*, 14, 1–5.
- Java, A., Song, X., Finin, T., & Tseng, B. (2007). *Why we twitter: understanding microblogging usage and communities*. Paper presented at the Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis.
- Karpp, D. (2012). Social science research methods in Internet time. *Information, Communication and Society*, 15(5), 639–661.
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006), 932.
- Kim, S. M., & Hovy, E. (2007). *Crystal: Analyzing predictive opinions on the web*. Paper presented at the Proceedings of the 2007 Joint Conference on Empirical Methods in

- Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), 1056–1064.
- King, I. (2010). *Introduction to social computing*. Paper presented at the Database Systems for Advanced Applications, 428–484.
- Kober, H., Barrett, L. F., Joseph, J., Bliss-Moreau, E., Lindquist, K., & Wager, T. D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. *Neuroimage*, 42(2), 998–1031.
- Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science*, 311(5757), 88–90.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: The dangers of double dipping. *Nature Neuroscience*, 12(5), 535–540.
- Laird, A. R., Lancaster, J. J., & Fox, P. T. (2005). Brainmap. *Neuroinformatics*, 3(1), 65–77.
- Landsall-Welfare, T., Lampos, V., & Cristianini, N. (2012). *Effects of the recession on public mood in the UK*. Paper presented at the Proceedings of the 21st international conference companion on World Wide Web, 1221–1226.
- Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., ... Alstyn, M. (2009). Computational social science. *Science*, 323(5915), 721–723.
- Lieberman, M. (2006). Sex-linked lexical budget. Retrieved October 26, 2013, from <http://itre.cis.upenn.edu/~my/languageelog/archives/003420html>.
- Mao, H., Counts, S., & Bollen, J. (2011). *Computational economic and finance gauges: Polls, search, and twitter*. Paper presented at the Meeting of the National Bureau of Economic Research-Behavioral Finance Meeting, Stanford, CT.
- Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, and Computers*, 33(4), 517–523.
- Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slater, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science*, 317(5834), 82–82.
- Mihalcea, R., & Liu, H. (2006). *A corpus-based approach to finding happiness*. Paper presented at the AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, 139–144.
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7(3), 221–237.
- Mishne, G., & Glance, N. S. (2006). *Predicting movie sales from blogger sentiment*. Paper presented at the AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, 155–158.
- O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). *From tweets to polls: Linking text sentiment to public opinion time series*. Paper presented at Proceedings of the International AAAI Conference on Weblogs and Social Media, Washington, DC.
- Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage*, 45(1), S199–S209.
- Poldrack, R. A., Kittur, A., Kalar, D., Miller, E., Seppa, C., Gil, Y., ... Bilder, R. M. (2011). The cognitive atlas: Toward a knowledge foundation for cognitive neuroscience. *Frontiers in Neuroinformatics*, 5, 17.
- Quercia, D., Ellis, J., Capra, L., & Crowcroft, J. (2011). *In the mood for being influential on Twitter*. Paper Presented at the Privacy, Security, Risk and Trust (Passat), 2011 IEEE 3th International Conference on and 2011 IEEE 3rd International Conference on Social Computing (socialcom), 307–314.
- Quercia, D., Kosinski, M., Stillwell, D., & Crowcroft, J. (2011). *Our Twitter profiles, our selves: Predicting personality with Twitter*. Paper presented at the Privacy, security, risk and trust (passat), 2011 IEEE 3th International Conference on and 2011 IEEE 3rd International Conference on Social Computing (socialcom), 180–185.
- Schuler, D. (1994). Social computing. *Communications of the ACM*, 37(1), 28–29.
- Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, 11(1), 87–104.
- Strapparava, C., & Valitutti, A. (2004). Wordnet-Affect: An affective extension of WordNet. In *Proceedings of the 4th international conference on language resources and evaluation* (pp. 1083–1086). Lisbon, Portugal.
- Suri, S., & Watts, D. J. (2011). Cooperation and contagion in web-based, networked public goods experiments. *PloS One*, 6(3), e16836.
- Szell, M., Lambiotte, R., & Thurner, S. (2010). Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences of the United States of America*, 107(31), 13636–13641.
- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95(5), 1202–1216.
- Wager, T. D., Barrett, L. F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., ... Mize, J. (2008). The neuroimaging of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 249–271). New York: Guilford.

- Wilson, T., Wiebe, J., & Hoffman, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In: *Proceeding HLT '05 Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing* (pp. 347–354). Vancouver, B. C., Canada.
- Yang, J., & Counts, S. (2010). Predicting the speed, scale, and range of information diffusion in Twitter. *ICWSM, 10*, 355–358.
- Yarkoni, T. (2010a). The abbreviation of personality, or how to measure 200 personality scales with 200 items. *Journal of Research in Personality, 44*(2), 180–198.
- Yarkoni, T. (2010b). Personality in 100,000 words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44*(3), 363–373.
- Yarkoni, T. (2012). Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science, 21*(6), 391–397.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*(8), 665–670.
- Yarkoni, T., Poldrack, R. A., Van Essen, D. C., & Wager, T. D. (2010). Cognitive neuroscience 2.0: Building a cumulative science of human brain function. *Trends in Cognitive Sciences, 14*(11), 489–496.
- Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter “I hope it is not as bad as I fear”. *Procedia-Social and Behavioral Sciences, 26*, 55–62.

## Psychoinformatics: The New Development of Psychology in the New Era

XUE Ting<sup>1,2</sup>; CHEN Hao<sup>3</sup>; LAI Kaisheng<sup>3</sup>; DONG Yinghong<sup>4</sup>; YUE Guo-an<sup>3</sup>

(<sup>1</sup> College of Humanities and Management, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China)

(<sup>2</sup> Department of Sociology, Faculty of Social Sciences, VU University Amsterdam, Amsterdam 1081HV, the Netherlands)

(<sup>3</sup> Department of Social Psychology, Nankai University, Tianjin 300071, China)

(<sup>4</sup> Department of Psychology, School of Educational Science, Ludong University, Yantai 264011, China)

**Abstract:** Psychoinformatics, a product of the information era, is an inter-disciplinary research area that incorporates computational techniques and information sciences to study human mind. It innovated how scholars acquire, organize, interpret and synthesize traditional psychological data. With its advantage to handle large dataset it shows great potential in exploring psychological mechanisms and behavioral patterns at both individual and group levels. Psychology researchers are able to inspect existing psychological theories and hypotheses from a novel perspective. The current review aims to provide guidance for readers to witness the development of this emerging psychological field. Future challenges will be discussed as well.

**Key words:** Psychoinformatics; psychology; information sciences; computer techniques