工程科学学报,第47卷,第9期:1841-1852,2025年9月

Chinese Journal of Engineering, Vol. 47, No. 9: 1841–1852, September 2025

https://doi.org/10.13374/j.issn2095-9389.2025.03.17.001; http://cje.ustb.edu.cn

# 基于情绪关联的多模态认知负荷量化

万 菲, 刘子阳, 支瑞聪™

北京科技大学计算机与通信工程学院,北京 100083 図通信作者, E-mail: zhirc research@126.com

摘 要 现代人机交互系统信息高度密集、任务复杂多变. 情绪是交互任务中的固有心理产物, 为探讨人机交互任务中操作员脑力负荷的变化, 依据认知负荷理论和成因分析, 将认知负荷划分为记忆力负荷、专注力负荷、情绪情感, 分别对应内部认知负荷、外部认知负荷、关联认知负荷. 以情绪作为内部认知负荷与外部认知负荷联系的桥梁, 继而对整体人机交互的认知负荷进行分析. 同时基于多模态数据协同, 提出了人机交互过程中操作人员脑力负荷动态预测模型. 为证明并校验所提出的模型, 选取 53 名被试完成 3 种渐进式难度的心理学范式. 结果验证了不同类型负荷激发任务下多模态信号、不同等级负荷激发任务下多模态信号与被试报告的主观评价的主效应显著(P<0.05); 总脑力负荷预测值与被试主观评价、子负荷预测值与子负荷量表主观评价呈显著正相关; 平均脑力负荷预测值与对应多模态信号呈显著正相关. 所提模型对人机交互过程中操作人员脑力负荷的动态预测与评价具有应用价值.

关键词 人机交互系统;认知负荷;负荷量化;多模态信号;负荷预测

分类号 TG142.71

# Emotion correlation-based multimodal cognitive load quantification

WAN Fei, LIU Ziyang, ZHI Ruicong<sup>™</sup>

School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China Corresponding author, E-mail: zhirc\_research@126.com

ABSTRACT Modern human—computer interaction (HCI) systems are characterized by high information density and complex tasks. In such systems, information is often presented in large volumes at rapid speeds, requiring users to process and comprehend it quickly. At the same time, tasks are intricate and multifaceted, involving multiple steps and decision-making processes. Monitoring the cognitive load in HCI tasks is of paramount importance because it can significantly enhance system performance and reduce the likelihood of operational errors. When users experience excessive cognitive load, their ability to process information and make decisions may be impaired, leading to decreased efficiency, increased error rates, and even system failures. By effectively monitoring cognitive load, system designers can identify potential bottlenecks and optimize the user experience to ensure smooth and efficient interaction. Cognitive load theory, supported by extensive causal research, provides a framework for understanding and categorizing cognitive load into three distinct types: internal, external, and associative. Internal cognitive load is closely tied to the brain's information processing mechanisms, particularly memory resources. It reflects the mental effort required to encode, store, and retrieve information from memory. Conversely, external cognitive load stems from the presentation and difficulty of learning materials. It is influenced by how information is organized and displayed to the user, as well as the inherent complexity of the task. Both internal and external cognitive loads are critical factors that affect a user's cognitive state during HCI tasks. However, directly measuring these two types of cognitive load is often challenging and time-consuming because it requires invasive or complex experimental setups. Given the limitations of direct measurement, this study proposes an innovative approach by leveraging the relationship between negative emotions and

收稿日期: 2025-03-17

基金项目: 中国实验室创新基金资助项目(2022-JCJQ-LA-001-080)

multimodal signals as a proxy for direct cognitive load measurement. Emotions, especially negative ones, are inherent psychological products of interactive tasks and are closely linked to cognitive processes. When users experience a high cognitive load, they often exhibit emotional responses such as frustration, anxiety, or stress. These emotional states can be reflected in various physiological and behavioral signals such as facial expressions, voice tone, eye movements, and physiological indicators such as heart rate and skin conductance. By capturing and analyzing these multimodal signals, we can indirectly infer the cognitive load experienced by users. To simulate the occurrence of cognitive load in HCI tasks, three repeatable visual stimuli were designed: digit memory, tracking, and combined memory-tracking tasks. The digit memory task focuses on memory by requiring participants to remember sequences of numbers. The tracking task emphasizes attention by asking participants to follow a moving target on the screen. The combined memorytracking task integrates both memory and attention by requiring participants to remember numbers while simultaneously tracking a moving target. The difficulty of each task type increases progressively, and the overall stimulus difficulty was modeled as a linear increase. This design allowed for a systematic investigation of how cognitive load evolves as task complexity increases. A total of 53 participants completed the designed paradigm experiments. They were required to complete the NASA-TLX workload assessment and PANAS emotion assessment scales. Several multimodal and multichannel signals during the experiment were recorded for subsequent analysis, including physiological data, eye-tracking data, and behavioral responses. This study proposes a novel cognitive load measurement method to quantify the psychological stress generated by subjects during HCI. By analyzing the differences in emotionrelated indices, conducting quantitative calculations, and performing classification verification, the study aims to establish a robust model for cognitive load assessment. The results demonstrated that the proposed progressive paradigm successfully induced a psychological load in the subjects. A significant correlation was found between the NASA-TLX scores and PANAS negative emotion scores, indicating that as the cognitive load increased, participants tended to experience more negative emotions. During the experiment, numerous multimodal and multichannel signals exhibited significant differences, suggesting that emotion-related signals undergo marked changes during cognitive load induction. According to the proposed calculation method, a significant correlation exists between the quantitative load and subjective load-level reports of the subjects. This finding validates the effectiveness of the proposed method in capturing cognitive load variations. Using different signals, the linear classification model achieved an accuracy of over 93%, which demonstrates its potential to accurately predict and assess cognitive load in real time. The ability to dynamically track the cognitive load can provide valuable feedback to system designers, enabling them to make timely adjustments and optimizations to the user interface and task design. Overall, the proposed model holds great promise for enhancing the user experience in HCI by providing a noninvasive and efficient way to monitor cognitive load.

KEY WORDS human-computer interaction system; cognitive load; load quantification; multimode signal; load predicting

随着计算机技术的飞速进步以及物联网的蓬勃兴起,人机交互(Human-computer interaction, HCI)中人的角色发生了深刻转变,不再仅仅是简单的操作执行者,更承担起了交互状态的监督重任,这种复合角色对操作者提出了极其严苛的信息处理要求.操作者一方面要精准执行指令,另一方面需实时监控交互系统的运行状态,对信息进行快速筛选、整合与反馈,使得其认知负荷呈持续上升态势.

在这样的背景下,分析认知负荷(Cognitive load)的重要性逐渐被学术界和实践领域广泛关注.通过深入探究人机交互过程中认知负荷的产生机制、影响因素及其变化规律,能够为优化交互系统设计提供关键依据.情感负荷理论□指出,在认知操作过程中,情感并非可有可无的附属品,而是一种具有消极属性的不确定性情绪,且这种情绪的强度会随着个体所感知到的时间压力的增大而呈现上升趋势.将该理论置于认知负荷理论的框架

之下进行审视,情绪加工这一重要环节可被重新诠释为情绪如何对工作资源的分配产生影响的问题.

基于上述理论认知,本研究将情感创新性地构建为一种内部负荷与外部负荷的关联体系,通过有效的分离不同类型的认知负荷并考察他们与不同生理、行为、视频信号之间的关联.具体而言,通过同步采集并观察与消极情绪相关的多路信号以及自我报告的情绪体验等数据,综合运用统计分析、机器学习算法等前沿技术手段,构建起一套认知负荷量化评估体系,从而为深入探究人机交互过程中个体的认知与情感交互作用提供有力的理论支撑与实践工具,进一步推动人机交互领域的精细化、智能化发展.

#### 1 情绪与认知负荷

认知心理学框架下,人类认知过程本质上体现为心理图式的动态更新与系统性扩展<sup>[2-3]</sup>.这一

理论视角将认知负荷解构为三个维度:涉及信息 呈现方式的外在认知负荷(External cognitive load), 表征知识本质属性的内在认知负荷(Internal cognitive load)、以及反映认知资源优化配置的关联认 知负荷(Associative cognitive load). 基于 Sweller 等<sup>[2]</sup> 的奠基性研究,总认知负荷被量化为三者的代数和(如图 1 所示),这为理解学习过程的认知机制提供了量化模型.

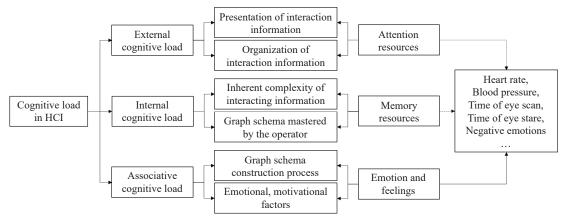


图1 人机交互过程中认知负荷的成因分析与多模态行为信号

Fig.1 Analysis of the causes of cognitive load during human-computer interaction

外在认知负荷作为"无效负荷"的典型表征, 其形成机制主要与教学设计的结构性缺陷相关. 当知识呈现方式违背认知规律(如多媒体信息不 同步、图文编排失当)或存在干扰要素(如冗余视 觉刺激、无关听觉信息)时,将迫使认知系统进行 额外的信息筛选与整合操作,从而产生与核心学 习目标无直接关联的认知消耗[4]. 针对脑部的红外 成像研究显示,认知负荷的激活区域主要集中于 背外侧前额叶皮层,该区域负责执行注意资源的 分配管理,这从神经机制层面解释了外在负荷对 注意系统的特异性影响[5]. 当学习者面临知识储备 不足与信息加工需求过载的双重挑战时,工作记 忆系统将承受的压力将呈现指数级增长[4,6]. 这种 现象的神经基础可追溯至前额叶皮层与负责记忆 功能的海马体的协同工作机制——当新输入信息 与既存工作记忆信息缺乏有效联结时,神经元网 络的突触重构需要消耗大量认知资源以重新构建 新的记忆数据[7]. 关联认知负荷的产生机制则体现 了认知系统的适应性特征,是在学习过程中由于 外部负荷和内部负荷不匹配的情况下的固有产物. 从生物学的角度来说,其来源主要有两部分:大脑 皮质层的记忆加工转换[8] 与学习者对注意力资源 的整合使用[9].

认知控制与情绪认知理论认为,情绪系统可以通过前额叶皮层与顶叶皮层的后感觉区域的动态耦合,通过工作记忆与自上而下的注意力行为,从而持续调节着工作记忆中记忆资源(对应内在

认知负荷)与注意资源(对应外在认知负荷),这也在神经影像学研究中得到了生物学证据<sup>[10]</sup>.因此,情绪并非仅仅是认知活动的伴随现象,而是构成认知过程的结构性要素,情绪机制因此被重新定位为认知资源分配的核心调节器<sup>[1,11]</sup>.在此前提下,情感能够作为衡量用户感知负荷是否超载的指标,并能够针对用户负荷节点进行对应认知结构的调整,从而优化整体认知路径<sup>[12]</sup>.

认知功能与情绪功能的交互作用构成动态的 双向调节系统,这一现象在神经教育学领域被称 为"认知-情绪耦合效应"[13]. 从认知功能影响情绪 的角度而言,由于相关认知负荷来自学习者将长 期记忆的相关信息与来自当前学习任务的信息相 关联的过程,情绪在此时属于分配给处理内在认 知负荷的工作记忆资源和处理外在认知负荷的注 意力资源.一方面,情绪能够影响内在认知负荷: 情绪通过将长期记忆中的相关信息投入到工作记 忆中,可以作为记忆资源的有效提取线索.此外, 情绪一致性理论认为情绪一致性使认知过程变得 容易. 通常来说, 积极的信息在积极的情绪下更容 易回忆,而消极的信息更容易在消极的情绪下回 忆. 另一方面, 情绪能够影响外在认知负荷: 情绪 作为一种线索促进认知加工,注意力假设认为情 绪通过分配注意力资源来影响学习过程中的认知 加工. 通常来说, 积极的情绪会扩大注意力的范 围,而消极的情绪会导致注意力的集中缩小.这一 点得到了眼动模式分析的支持[14].

传统量表通常将认知负荷视为单一维度进行 整体评估,难以区分内在、外在和相关认知负荷的 独立作用.此外,传统量表多基于被试的自我报告, 容易受到主观偏差的影响. 而随着各种传感器技术 的成熟、应用成本的降低,越来越多的研究者尝试 使用物理信号来评估认知负荷[15]. 此外, 考虑到情 绪反应的复杂性与认知过程中的微表情的细微性, 基于多模态数据的评估方法具有更高的鲁棒性和 抗噪声能力[16]. 目前,诸如生理信号(包括心率、血 压等)、行为信号(如眼动追踪)、视频信号(如面部 肌肉运动单元)等单模态数据已被用于认知负荷评 估,且因不同模态数据之间可以实现信息互补,故 能得到更准确的评估结果[17-18]. 如图 1 所示, 受到 消极情绪影响,除了常见的消极面部表情外,人的 心率、血压,乃至眼动信号都会出现相应的变化.而 这些信号也同时能够反应消耗注意力资源和记忆 资源的外在认知负荷与内在认知负荷的情况.

本研究结合人机交互过程中认知负荷的成因分析,从基于情绪产生的关联负荷的角度出发<,利用所提出的人机交互模拟任务范式刺激,通过NASA-TLX负荷量表、PANAS情绪情感量表的被试主观评价,采集多模态行为信号,验证了认知脑力负荷动态预测模型.

#### 2 国内外研究现状

传统的人机交互效能的评估大多采用专家打分、问卷调查等方法,评估结果主观性强、缺乏科学依据,且没有统一的衡量标准. Lorenz 等[19] 通过主任务测试、NASA-TLX、SUS 量表问卷调查等方法对挖掘机控制系统人机交互效能进行评估. Liu 等[20] 将指控系统人机交互效能划分为 126 个可用性指标对人机交互效能进行评估. 然而,这些方法缺乏统一科学依据,并且具有主观性强、没有统一的衡量标准的问题. 相较于传统的主观评估方式,观测指标测量法具有采集数据科学、客观,计算结果可信度高等优点.

Trisitichoke 等<sup>[21]</sup> 发现面部表情与感知可理解性高度相关,这表明面部表情可用来评估与语言理解有关的认知负荷. 眼动追踪技术可通过采集学习过程中的注视、扫视、瞳孔变化等眼动行为数据来评估认知负荷, Joseph 和 Murugesh <sup>[22]</sup> 研究了人机交互中产生的认知负荷与眼动指标间的关系, 薛耀锋和李卓玮<sup>[23]</sup> 结合注视时间、注视点个数、瞳孔直径等指标构建了基于眼动追踪技术的在线学习认知负荷量化模型, 并通过实验验证了

量化模型的可行性和准确性.认知负荷与大脑紧密相关,且脑电信号与认知负荷具有很高的同步性<sup>[24]</sup>,因此使用脑电信号来评估认知负荷也成为一个常见的研究思路.例如,Li等<sup>[25]</sup>通过提取脑电信号不同频段上的功率谱密度特征,对模拟驾驶者的认知负荷进行了评估.随着各种传感器技术的成熟和应用成本的降低,使用多模态数据来评估认知负荷受到更多研究者的欢迎<sup>[26-27]</sup>.通过多特征融合方法,结合 BP 神经网络和 DS 证据理论分析提高识别精度.

一些研究从知觉负荷和认知负荷的角度探讨了工作记忆、注意资源与负性情绪情感之间的关系<sup>[28-30]</sup>,并验证了消极情绪是影响注意功能和工作记忆的重要因素,提出了认知呼和环境下,情绪体验与工作记忆过程和注意力功能存在的交互作用.然而,这些研究并没有将三者作为一个整体对象,而是分别探讨情绪与注意资源、情绪与工作记忆之间的关联.

为了研究多模态行为分析和认知能力指标体系,搭建视觉感知、物理感知、生理信号等多通道信息采集系统,本文利用认知功能与情绪功能间的双向关系,结合情绪的外在表现模态和人类认知模式,开展多路信号感知协同优化的情绪智能测量方法研究,实现生理感知和心理认知之间的关联评估.

#### 3 研究方法

参考多模态数据的相关应用研究,从中汲取实验设计、数据处理等方面的经验,本研究设计了基于多模态数据的在线学习认知负荷评估研究框架.该框架包含多模态数据采集、多模态特征提取、评估模型构建三个部分.

## 3.1 认知负荷行为数据采集

我们将从三个部分分别介绍认知负荷行为数据采集,包括实验被试及准备工作、心理实验范式设计和多通道信息采集系统.

#### 3.1.1 实验被试及准备工作

本实验通过高校论坛发布招募通知,共召集了53名北京科技大学本科生和研究生参与实验,其中女性24位,男性29位,被试年龄在19~29岁之间(均值(Mean)=24.21,方差(SD)=3.29).所有被试均为右利手,视力或矫正视力达1.0,无色盲色弱等眼部疾病,无任何行为或认知障碍.每个被试均自愿参与本实验,在充分了解实验内容后签署知情同意书.实验结束后,被试可获取劳务报酬.

实验地点位于北京科技大学机电信息楼 416, 实验环境安静, 无明显干扰噪声. 实验所需硬件设备包括: 一台 16 GB 内存计算机、Dell 22 英寸显示屏和一套 Logitech 无线键鼠. 实验任务界面显示在

22 英寸显示器上,被试佩戴耳机(用于隔绝外部噪音),右手使用鼠标来执行自定义的符合刺激实验. 实验按照实验任务分为三个阶段,每个阶段都通过参数控制题目难度的渐进性,如图 2 所示.

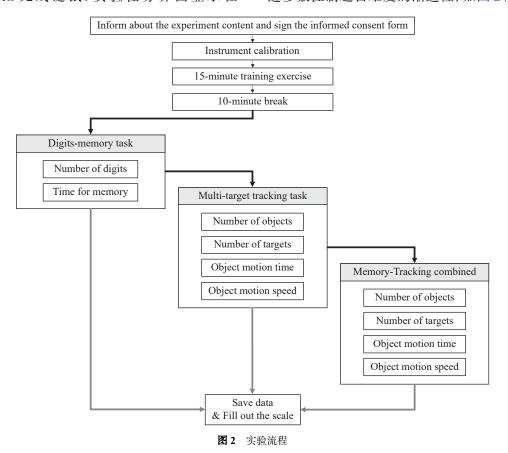


Fig.2 Experimental process

实验前一天要求被试进行充分的休息,不饮酒和熬夜.被试到达实验地点后,主试向其讲解实验内容和注意事项,并签署知情同意书.在开始正式实验前,被试首先进行 10 min 自定义预实验的培训和练习. 预实验的培训练习被设计是简单且低负荷的,以确保在正式实验中不会产生学习效应,即正式实验期间绩效不会随着实验的进行而出现显著提高的现象.实验的总持续时间约为 1.5 h,包括培训练习 (10 min)、低负荷阶段的正式实验任务 (20 min)、中等负荷阶段的正式实验任务 (20 min)、高负荷阶段的正式实验任务 (15 min)、仪器校准、问卷填写和休息时间.

#### 3.1.2 心理实验范式设计

普通人机界面的信息容量大、结构关系错综复杂,不同的操作者,尤其是未经训练过的普通人员,对这样的系统的感知负荷是不一致的.为了能够更全面且可控的获得被试者在完整的刺激时段内的情绪状态变化、专注力波动和认知负荷的诱发情况,需要设计单独且可重复的视觉刺激.我们设计以下的实验任务范式用于测量认知负荷,并记录相关的行为数据,这些实验任务将会按照顺序依次进行.总体来说,刺激的产生和被试作答时间按照图3分布.



Fig.3 Stimulus experiment time distribution

(1)数字记忆任务: 数字记忆任务来源自数字/字母的 N-back 实验<sup>[31]</sup>. N-back 实验要求被试判断

每一个出现的刺激与前面呈现过的第*n*个刺激是 否相匹配. 该实验可通过控制 *n* 的大小来操作工 作记忆的负荷,从而考察不同记忆负荷下工作记忆的加工机制.研究者也常采用 N-Back 任务来测量工作记忆容量.由于传统 N-back 实验中要求具备听觉刺激,受到实验器材和平台自我搭建的限制,我们将 N-back 实验修改为一个完整的数字记忆实验,如图 4 所示.首先,屏幕上会呈现一串位

置随机的连续数字,在一定的记忆时间过后,数字被黑色的圆圈掩盖.被试被要求依照记忆中数字所处于的位置,按照数字的大小顺序从小到大依次点击黑色的圆圈.被试必须按顺序正确且完整的复述整串数字,只要有一个数字的位置复述错误,则整道题目被判罚回答错误.此部分共70道题目.

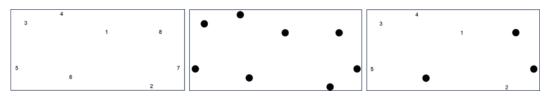


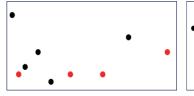
图 4 数字记忆任务示意图,刺激呈现期(左)、刺激运动期(中)、被试作答期(右)

Fig. 4 Diagram of the digital memory task, stimulus presentation period (left), stimulus movement period (middle), and subject response period (right)

(2)多目标追踪任务<sup>[32]</sup>:多目标追踪范式被广泛用于视觉认知的研究,主要用于在动态和并行信息的注意加工机制的探讨上.如图 5 所示,在此任务中,电脑屏幕上将出现多个客体(圆圈),其中一部分以一定的方式标识为目标,其余为非目标.之后所有客体开始随机运动,而被试需要对目标.进行追踪,并在运动停止后对目标进行全部报告.被试必须完整的汇报所有的被标识目标.此部分共73 道题目.

(3)混合记忆追踪任务:考虑到操作者的认知 负荷受到学习者的已有经验影响,为了使被试在 整个数据采集的过程中对题目保持"新鲜感",从 而能够持续的增加被试所感受到的认知负荷压 力,考察被试在连续的视觉刺激下的效率、反应时 间、决策等的水平,我们追加了混合记忆追踪任务,如图 6 所示. 屏幕上会出现具有连续标号的客体(圆圈),随后其中不连续标号的一部分被标识以提醒被试. 之后所有标号消失, 所有客体开始随机运动. 运动完成后,被试和多目标追踪任务中一样,需要对目标进行追踪, 不同的是, 这次被试需要按照一开始被标识客体的编号从小到大依次按顺序进行选择. 实验者被要求必须同时回答正确被标识的客体以及它们的编号顺序. 为了同时记录记忆负荷和专注负荷的情况,混合记忆追踪任务允许被试在作答编号顺序错误的情况下继续作答追踪目标. 此部分共49 道题目.

实验采用被试内设计,自变量为任务复杂性, 因变量为被试人员脑力负荷变化.对于自变量而





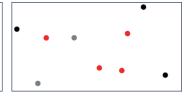


图 5 多目标追踪任务, 刺激呈现期(左)、刺激运动期(中)、被试作答期(右)(在被试作答期中, 答案实时判断, 作答正确显示为红色, 否则为灰色) Fig.5 Multi-objective tracking task, stimulus presentation period (left), stimulus movement period (middle), and subject response period (right) (during the test period, the answer is judged in real time, and the correct answer is shown in red; otherwise, it is gray)



图 6 混合记忆追踪任务,刺激呈现期(左)、刺激运动期(中)、被试作答期(右)(在被试作答期中,答案实时判断,作答正确显示为红色,否则为灰色)

**Fig.6** Mixed memory tracking task, stimulus presentation period (left), stimulus movement period (middle), and subject response period (right). (during the test period, the answer is judged in real time, and the correct answer is shown in red; otherwise, it is gray)

言,任务复杂性采取递进式设计,包含三个阶段:记忆负荷任务( $L_{\rm m}$ )、专注负荷任务( $L_{\rm a}$ )、混合记忆—专注负荷( $L_{\rm c}$ )任务,分别对应着实验任务中数字记忆任务( $L_{\rm N}$ )、多目标追踪任务( $L_{\rm VS}$ )和混合记忆追踪( $L_{\rm NVS}$ )任务.每种负荷激发任务通过参数控制(如表 1 和表 2 所示).对因变量而言,为了与被试的感受达成统一,被试需要在实验过程中感到负荷跃迁时做出按键反应.被试在实验进行中允

许进行多次按键,但被要求在实验结束后确定的选择最能代表自身负荷程度变化的三个节点.为了减轻被试填写量表的负担,实验任务完成后,被试仅被要求回忆负荷最高的状态,并填写 NASA-TLX 量表 (NASA Task load index)、情感 PANAS (Positive and negative affect schedule)量表.之后,我们按照低负荷=30%最高负荷,中等负荷=70%最高负荷进行处理.

表1 不同负荷激发任务的难度参数

Table 1 Difficulty parameters of different load excitation tasks

Task Name	Positive correlation difficulty parameter	Negative correlation difficulty parameter
Digits memory task	Digits number	Memory time
Multitarget tracking task	Object number, target number, motion time, motion speed	Memory time
Memory tracking task	Object number, target number, motion time, motion speed	Memory time

#### 表 2 符号对照表

Table 2 Symbol comparison table

Object	Target	Motion	Motion	Memory
Number	Number	Time	Speed	Time
$\frac{N_{\rm y}}{N_{\rm y}}$	N <sub>t</sub>	$T_{ m mov}$	V	$T_{\text{mem}}$

#### 3.1.3 心理实验范式难度量化

同时,我们提出了一种任务难度的量化方法, 三种任务的难度 $D_L$ 计算方法分别如下:

$$D_{L_{\rm N}} = N_{\rm t}/(T_{\rm mem} + \alpha)^{1/2} \tag{1}$$

$$D_{L_{VS}} = (N_{V} \times N_{t} \times T_{mov} \times V / (T_{mem} + \alpha)^{1/2})^{1/2}$$
 (2)

$$D_{L_{\text{NVS}}} = N_{\text{t}}/(T_{\text{mem}} + \alpha)^{1/2} + (N_{\text{y}} \times N_{\text{t}} \times T_{\text{mov}} \times V/(T_{\text{mem}} + \alpha)^{1/2})^{1/2}$$
(3)

其中, $\alpha$ 用于调节线性难度的拟合程度,其取值来自人眼视觉暂留效应,取值范围在  $0.1 \sim 0.4$ . 本实验中取值为 0.4. 之后,所有的难度系数使用 maxmin 方法进行标准化,从而得到近似线性上升的难度设置.整体实验难度系数的线性回归分析表如表 3 所示,其中 $R^2$ 为决定系数,代表了自变量对因变量的解释程度,P为显著性指标.

表3 所有刺激任务难度的线性回归分析

Table 3 Linear regression analysis of the difficulty of all stimulus tasks

$R^2$	P
0.902	$0.000^{*}$

Note: \* indicates the significance.

#### 3.1.4 多通道特征数据采集

在实验过程中,通过不同的设备记录与情绪 相关的多种行为数据

- (1)视频特征提取:视频数据通过 Open Face 逐帧分解图像,对每一帧图像进行人脸检测、面部特征点估计、面部特征输出等步骤从而计算抽象特征,集成后获得包含眼部视线方向、头部姿态、面部动作单元等特征向量,并通过统计分析计算相应的均值、标准差,共提取 34 个具体指标.
- (2)眼动追踪指标提取:眼动数据经过预处理,去除异常点后,计算出扫描路径长度,扫描路径时间,注视时间等8个眼动指标
- (3)生理信号特征提取:生理数据则先经过预处理去除异常点,使用巴特沃斯滤波器进行平滑,提取状态参数,变异参数,峰值参数,时间参数和区域参数共6类特征51个指标.

所有指标如表 4 所示.

# 3.2 认知负荷行为数据分析

#### 3.2.1 多路信号与量表数据处理方法

为了保证量表的可信度,我们利用 Cronbach's alpha 参数和 KMO(Kaiser-Meyer-Olkin) 指数在描述性统计中参数计算被试作答量表的信度与内容效度. 如果被试在实验前后作答的量表得分未能通过信度与效度检验,则作废该被试的量表回答.

我们使用工具 SPSS 27.0 进行差异性分析. 根据两两负荷标度的差异,采用单因素方差分析,使用 Levene test 进行方差齐性检验. 如果不满足方差齐性,使用 Brown-Forsythe 值替代方差齐性时使用的 F值,用于检验各组均值是否相等. 满足方差齐次时,使用 Bonferroni's multiple comparison test进行事后检验,不满足方差齐次时,使用 Dunnett's T3 进行事后检验. 我们输出所有具有显著差异的

表 4 多通道数据的相关指标

Table 4 Related indicators of multichannel data

Data Type	Indicator description	Feature encoding	
Blood	Status parameter	xx_max, xx_min, xx_mean, xx_median	
	Variant parameter	xx_range, xx_sd, xx_iqr, xx_idr, xx_mad	
pressure	Apex parameter	xx_tmax	
	Time parameter	xx_dgm, xx_dga	
	Area parameter	xx_area, xx_arear	
Heart rate	Standard deviation of normal sinus beat interval, standard deviation of differences of all adjacent NN intervals	hr_sdnn, hr_sdsd	
	Average power across different frequency bands	hr_tp, hr_lf, hr_hf, hr_lh, hr_hfp, hr_lfp	
Eye	Scan, focus path length and time	eye_scan_length, eye_scan_time, eye_stare_time, eye_stare_count	
Action unit (AU)	Mean intensity and standard deviation of AU01-AU45	AUxx_mean, AUxx_std	

信号,并利用支持向量机(Support vector machine, SVM)分类做验证.

为了进行数据拟合,我们取每份数据的均值. 针对标量数据,我们取变异值(与上一阶段的值的 差)作为研究对象.

#### 3.2.2 自变量和因变量

本实验的自变量为主任务复杂性.其中主任务复杂性包括3种设置,分别是数字记忆任务、多目标追踪任务和混合记忆追踪任务. 三种刺激的连续设置会导致累积负荷的不断升高. 在每种刺激设置中,通过相应的参数可以控制难度的不断递增. 本实验的因变量包括主任务绩效和主观脑力负荷,且绩效与脑力负荷之间存在函数关系.

自定义的刺激任务将每个子任务的事件发生、处理和被试的反应情况记录在主文件 (Master files) 中. 因此,主任务绩效可通过对主文件进行分析提取,包括被试的作答情况(所有点击和点击顺序)和反应时间. 主任务绩效数值越高,表明绩效越差. 由于主任务中三个子任务的绩效衡量单位不统一,因此通过 z-score 标准化后相加,以获得主任务的总体绩效分数. 所有子任务的均被标准化后,使用相同的权重分别计算完成时与正确率.

#### 3.2.3 认知负荷预测计算方法

心理负荷理论认为:认知负荷和个人任务表现呈倒U型关系(图7),即太高或太低的任务负荷均不能带来最佳的任务表现,中等强度的任务负

荷有利于个人发挥最佳的绩效水平.本文使用倒 U型理论来建模被试认知负荷与绩效之间的关 系,并使用答题准确率(acc)代替绩效计算结果.

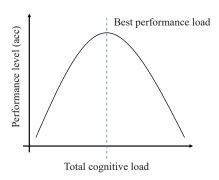


图 7 绩效水平和认知负荷之间的倒 U 型关系

Fig.7 Inverted U-shaped relationship between performance level and cognitive load level

假设负荷(load)与绩效的取值范围均为[0,1],最佳绩效负荷取值为 0.5,本文以作答准确率为绩效.负荷(load)-绩效(acc)倒 U型关系可以被简化为函数:

$$acc = f(load) = -4load^2 + 4load$$
 (4)

则被试认知负荷的计算方法为:

load = 
$$f^{-1}(\text{acc}) \times D_L$$
 =   

$$\begin{cases}
-(\sqrt{1 - \text{acc}} - 1)/2 \times D_L, D_L = D_{L_N} \text{ or } D_L = D_{L_{VS}} \\
(\sqrt{1 - \text{acc}} + 1)/2 \times D_L, D_L = D_{L_{NVS}}
\end{cases}$$
(5)

#### 4 结果与讨论

#### 4.1 量表描述性统计

我们保存了 53 名被试的量表数据, 以验证 60 min 的渐进式心理刺激任务用于疲劳和消极情绪诱发的有效性.

## 4.1.1 脑力疲劳诱发验证

考虑到量表容量过大,被试在实验开始前并没有被要求填写 NASA-TLX 负荷量表,而是使用斯坦福困倦量表进行代替. 因正态性检验表明该数据不满足正态分布,因此使用 Wilcoxon 秩和检验(Signed-rank test)的非参数检验方法对任务前后的斯坦福困倦量表数据进行分析. 结果显示,渐进式心理刺激任务后的斯坦福困倦量表评分(均值(Mean)=5.12,方差(SD)=0.7)显著高于任务前(Mean=2.10, SD=1.20)(P<0.001). 这表明 60 min 的渐进式心理刺激任务成功诱发了被试的脑力疲劳状态,使得被试的心理负荷水平显著升高.

#### 4.1.2 消极情绪诱发验证

正态性检验表明该 PANAS 消极情绪量表数

据不满足正态分布,因此使用 Wilcoxon 秩和检验 (Signed-rank test)的非参数检验方法对任务前后的 PANAS 消极情绪量表数据进行分析. 结果显示,渐进式心理刺激任务后的量表评分(Mean=2.01, SD=0.07)(P<0.001). 这表明 60 min 的渐进式心理刺激任务成功诱发了被试的消极情绪. 人机交互中的脑力认知负荷与交互过程中的消极情绪得分存在显著相关性. 因此在难以直接计算脑力负荷时,可以考虑使用消极情绪相关的指标作为代替.

# 4.1.3 负荷与消极情绪得分关联性分析

我们对 NASA-TLX 负荷量表得分与 PANAS 消极情绪量表得分进行相关性分析, 采用皮尔逊相关系数法, 结果如表 5 所示. 这说明在人机交互的过程中, 负荷累积的结果与消极情绪的增长有显著关联. 因此, 利用多模态多路信号, 通过判断消极情绪的强度来预测负荷水平是一种合理的方法.

表 5 负荷与消极情绪相关性分析

 Table 5
 Correlation analysis of load and negative emotion

$R^2$	P	
0.876	$0.000^{*}$	

Note: \* indicates the significance.

# 4.2 多路信号的差异性分析

基于被试按键反应反馈的负荷等级,我们使用单因素方差分析多路信号数据.使用 Levene

test 进行方差齐性检验. 如果不满足方差齐性,使用 Brown-Forsythe 值替代方差齐性时使用的 F 值,用 于检验各组均值是否相等. 满足方差齐次时,使用 Bonferroni's multiple comparison test 进行事后检验, 不满足方差齐次时,使用 Dunnett's T3 进行事后检验.

表 6 的结果可知三方面内容.(1)在认知负荷累积的过程中,心率和血压在峰值和持续时间两个类别的参数上具有比其他类型的生理信号更显著的差异.这与消极情绪累加时,人们会出现呼吸加促、心跳加快等现象吻合.具有显著性差异的区域参数进一步证明了这一点.(2)眼动信息中,扫描和注视的时间明显加长,次数也明显增加的信息表明被试者在答题过程中注意力存在下降的情况.这与消极情绪带来的视觉粘滞效应和回视行为特征表征一致,也同样说明了负性刺激中的微眼动抑制机制带来的负性记忆固化.(3)对于面部动作单元,具有显著性差异的几项指标则均与消极情绪所对应的面部动作单元有很高的重复.

通过方差分析筛选出在不同负荷水平下差异均显著的特征,说明这些特征与负荷等级之间存在相关关系.进一步可以利用这些特征进行负荷预测.

#### 4.3 负荷预测

#### 4.3.1 阶段性显著差异指标分析

为了更好的进行分类分析,我们对三个负荷 阶段上的消极情绪相关的指标进行两两的单因素 方差分析.通过与表6的比较可以发现,在两两比

表 6 差异性信号分析

 Table 6
 Differential signal analysis

F-value	P	Feature	F-value	P	Feature	F-value	P
22.498	0.000*	hr_mad	33.151	0.000*	hr_mad	33.151	0.000*
105.792	$0.000^{*}$	hr_hfp	6.179	$0.000^{*}$	hr_lf	233.843	$0.000^{*}$
107.403	0.000*	hr_lmp	6.177	0.000*	hr_wpe	369.324	0.000*
44.910	0.000*	hr_tp	81.324	0.000*	hr_range	17.222	0.000*
78.240	0.000*	sbp_max	13.326	0.000*	sbp_arear	38.974	0.000*
44.546	0.000*	sbp_range	55.191	0.000*	sbp_sd	57.886	0.000*
36.248	0.000*	sbp_mad	167.771	0.000*	sbp_idr	65.308	0.000*
107.815	0.000*	dbp_mad	154.448	0.000*	dbp_range	5.296	0.000*
22.985	0.000*	dbp_arear	22.444	0.000*	dbp_sd	5.542	0.000*
46.698	0.000*	eye_stare_time	303.912	0.000*	eye_stare_count	124.731	0.000*
170.393	0.000*	eye_scan_time	124.562	0.000*	gaze_x_std	1.734	0.000*
12.809	0.000*	AU05_mean	38.642	0.000*	AU09_mean	47.480	0.000*
52.979	0.000*	AU05_std	21.151	0.000*	AU06_std	65.343	0.000*
82.920	0.000*	AU17_std	58.246	0.000*	AU20_std	99.931	0.000*
40.205	0.000*	AU26_std	46.239	0.000*			
	22.498 105.792 107.403 44.910 78.240 44.546 36.248 107.815 22.985 46.698 170.393 12.809 52.979 82.920	22.498       0.000*         105.792       0.000*         107.403       0.000*         44.910       0.000*         78.240       0.000*         44.546       0.000*         36.248       0.000*         107.815       0.000*         22.985       0.000*         46.698       0.000*         170.393       0.000*         12.809       0.000*         52.979       0.000*         82.920       0.000*	22.498         0.000*         hr_mad           105.792         0.000*         hr_hfp           107.403         0.000*         hr_lmp           44.910         0.000*         hr_tp           78.240         0.000*         sbp_max           44.546         0.000*         sbp_mad           107.815         0.000*         dbp_mad           22.985         0.000*         dbp_arear           46.698         0.000*         eye_stare_time           170.393         0.000*         AU05_mean           52.979         0.000*         AU05_std           82.920         0.000*         AU17_std	22.498         0.000*         hr_mad         33.151           105.792         0.000*         hr_hfp         6.179           107.403         0.000*         hr_lmp         6.177           44.910         0.000*         hr_tp         81.324           78.240         0.000*         sbp_max         13.326           44.546         0.000*         sbp_range         55.191           36.248         0.000*         sbp_mad         167.771           107.815         0.000*         dbp_mad         154.448           22.985         0.000*         dbp_arear         22.444           46.698         0.000*         eye_stare_time         303.912           170.393         0.000*         AU05_mean         38.642           52.979         0.000*         AU05_std         21.151           82.920         0.000*         AU17_std         58.246	22.498         0.000*         hr_mad         33.151         0.000*           105.792         0.000*         hr_hfp         6.179         0.000*           107.403         0.000*         hr_lmp         6.177         0.000*           44.910         0.000*         hr_tp         81.324         0.000*           78.240         0.000*         sbp_max         13.326         0.000*           44.546         0.000*         sbp_range         55.191         0.000*           36.248         0.000*         sbp_mad         167.771         0.000*           107.815         0.000*         dbp_mad         154.448         0.000*           22.985         0.000*         dbp_arear         22.444         0.000*           46.698         0.000*         eye_stare_time         303.912         0.000*           170.393         0.000*         eye_scan_time         124.562         0.000*           12.809         0.000*         AU05_mean         38.642         0.000*           52.979         0.000*         AU05_std         21.151         0.000*           82.920         0.000*         AU17_std         58.246         0.000*	22.498         0.000*         hr_mad         33.151         0.000*         hr_mad           105.792         0.000*         hr_hfp         6.179         0.000*         hr_lf           107.403         0.000*         hr_lmp         6.177         0.000*         hr_wpe           44.910         0.000*         hr_tp         81.324         0.000*         hr_range           78.240         0.000*         sbp_max         13.326         0.000*         sbp_arear           44.546         0.000*         sbp_range         55.191         0.000*         sbp_sd           36.248         0.000*         sbp_mad         167.771         0.000*         sbp_idr           107.815         0.000*         dbp_mad         154.448         0.000*         dbp_range           22.985         0.000*         dbp_arear         22.444         0.000*         dbp_sd           46.698         0.000*         eye_stare_time         303.912         0.000*         eye_stare_count           170.393         0.000*         AU05_mean         38.642         0.000*         AU09_mean           52.979         0.000*         AU05_mean         38.642         0.000*         AU06_std           82.920	22.498         0.000*         hr_mad         33.151         0.000*         hr_mad         33.151           105.792         0.000*         hr_hfp         6.179         0.000*         hr_lf         233.843           107.403         0.000*         hr_lmp         6.177         0.000*         hr_wpe         369.324           44.910         0.000*         hr_tp         81.324         0.000*         hr_range         17.222           78.240         0.000*         sbp_max         13.326         0.000*         sbp_arear         38.974           44.546         0.000*         sbp_range         55.191         0.000*         sbp_sd         57.886           36.248         0.000*         sbp_mad         167.771         0.000*         sbp_idr         65.308           107.815         0.000*         dbp_mad         154.448         0.000*         dbp_range         5.296           22.985         0.000*         dbp_arear         22.444         0.000*         dbp_sd         5.542           46.698         0.000*         eye_stare_time         303.912         0.000*         gaze_x_std         1.734           12.809         0.000*         AU05_mean         38.642         0.000* <td< td=""></td<>

较时,低负荷-中负荷的显著性差异指标与中负荷 -高负荷的显著性差异指标有所不同,而低负荷 -高负荷所具有的显著性差异指标则是低负荷-中 负荷/中负荷-高负荷所具有的差异指标的并集.我 们推测这有可能是不同指标的敏感性差异不同所导致的. 我们将这些在某个阶段更具有差异性的指标汇总在表 7 中. 其中, 在中负荷-高负荷具有更显著差异的指标用下划线表示.

表 7 阶段性显著差异指标

Table 7 Stage-wise significant difference indicators

Load level	Indicators					
Load level	Low load	Medium load	High load			
Low load	_	hr_wpe, sbp_max, sbp_range, sbp_sd, sbp_iqr, sbp_idr, sbp_mad, dbp_iqr, dbp_mad, eye_scan_length, eye_scan_time, AU9_mean, AU5_std, AU17_std, AU23_std, AU45_std	hr_wpe, hr_range, hr_mad, hr_tmax, hr_hfp, hr_lfp, sbp_max, sbp_range, sbp_sd, sbp_iqr, sbp_idr, sbp_mad, dbp_iqr, dbp_mad, dbp_range, dbp_sd, eye_scan_length, eye_scan_time, AU9_mean, AU5_std, AU17_std, AU23_std, AU45_std			
Medium load	hr_wpe, sbp_max, sbp_range, sbp_sd, sbp_iqr, sbp_idr, sbp_mad, dbp_iqr, dbp_mad, eye_scan_length, eye_scan_time, AU9_mean, AU5_std, AU17_std, AU23_std, AU45_std	_	hr_range, hr_mad, hr_tmax, hr_hfp, hr_lfp, dbp_range, dbp_sd			
High load	hr_wpe, hr_range, hr_mad, hr_tmax, hr_hfp, hr_lfp, sbp_max, sbp_range, sbp_sd, sbp_iqr, sbp_idr, sbp_mad, dbp_iqr, dbp_mad, dbp_range, dbp_sd, eye_scan_length, eye_scan_time, AU9_mean, AU5_std, AU17_std, AU23_std, AU45_std	hr_range, hr_mad, hr_tmax, hr_hfp, hr_lfp, dbp_range, dbp_sd	<del>-</del>			

## 4.3.2 负荷量化

根据 3.2.3 节提出的认知负荷预测测量方法, 我们利用表 6 的情绪相关指标计算了定量的认知 负荷强度,结果如图 8 所示,其中横轴时间(t)表示 当前被试作答的题号(为了显示方便,我们只截取 其中 5 人的认知负荷定量计算结果作为展示).

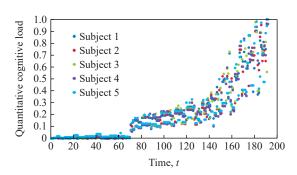


图 8 随时间变化的感知负荷定量度量情况

Fig.8 Quantitative measures of perceived load over time

值得注意的是,图 8 的结果显示定量计算出的 认知负荷在时间轴显示 t=70 和 t=150 时,被试者 恰巧分别进入第二阶段和第三阶段的交互任务, 此时计算出的定量负荷也出现了明显的跃迁现象. 我们使用三次方函数拟合计算出的定量认知负荷与被试主观反馈的认知负荷等级并进行回归分析,结果如表 8 所示.

表8 量化负荷与等级负荷间的三次方回归分析

Table 8 Cubic regression analysis between quantified load and grade load

$R^2$	F-value	P
0.980	4618.454	0.000*

由表 8 可知,使用三次方函数拟合的结果调整后的  $R^2 = 0.980$ , P < 0.001,这说明模拟的拟合程度较好,且回归结果具有显著性.

# 4.3.3 负荷分类预测

我们分别利用表 6 所列的指标和表 7 所列的指标针对所有负荷等级数据进行 5 折 SVM 分类,结果分别如表 9 和表 10 如所示.其中 acc 代表准确率, recall 代表召回率, fl 表示 fl 分数.

从表 9 和表 10 的结果可知, 使用针对情绪的 阶段性显著差异指标的结果要显著高于全部差异 性指标进行负荷等级分类的结果. 使用前者的预

表9 使用全部差异性指标的负荷等级分类结果

Table 9 Load classification results using all the difference indicators

Load level	acc/%	recall/%	f1/%
Low load	60.21	57.67	58.91
Mediumload	43.03	77.54	55.34
High load	100.00	0.70	1.39
overall	49.39	45.31	38.55

表 10 使用阶段性显著差异指标的负荷等级分类结果

Table 10 Load class classification results using stage significant difference indicators

Load level	acc/%	recall/%	f1/%
Low load	92.21	93.81	93.01
Medium load	98.39	98.13	98.26
High load	92.83	90.88	91.84
overall	94.54	94.27	94.37

测分类在全部类别的准确率上超过了 94.54%, recall 值达到了 94.27%, fl 值则为 94.37%.

这说明,不同等级负荷激发任务下多模态信号与负荷等级之间存在差异性.这可能是由于指标对于不同等级的负荷的敏感度不同,即一些指标可能对低负荷区域更敏感,而一些指标可能对高负荷区域更敏感.

# 4.3.4 适用场景讨论

根据已有的实验结果,人机交互的背景下,操作者的负面情绪与其感受到的认知负荷存在高度一致性(见 4.1 节),这为使用情绪测量作为间接评估脑力负荷的方法提供了理论依据.同时,多项生理指标、行为指标均显示出了在不同负荷等级下的差异化表现(见 4.2 节),这为所提出的间接测量方法提供了技术可行性.此外,根据表7的结果显示,不同的负性情绪指标对低负荷-中负荷以及中负荷-高负荷的敏感度不同,这也为不同的负荷测量场景提供了更精细的监测指标选择,因此在适用场景上具有一定的普适性.

总体而言,考虑到传统量表的局限性,多模态信号分析的必要性,以及行为、生理信号测量与消极情绪的关联性,情绪测量可以作为有效的替代或补充方法,尤其是在那些生态效度要求高与情感激发型的场景,例如模拟驾驶与课堂学习等任务.然而,受到本次实验被试的召集限制,目前仅能确定这种基于情绪的多模态认知负荷量化方法适用于青年群体.考虑到神经可塑性模型的动态适配性与年龄相关,超过此年龄范围的人群是否

适配这种间接测量方法仍需要进一步研究.

# 5 结论

- (1)人机交互中的脑力认知负荷与交互过程中的消极情绪得分存在显著相关性. 因此在难以直接计算脑力负荷时,可以考虑使用消极情绪相关的指标作为代替.
- (2)不同等级负荷激发任务下多模态信号与负荷等级之间存在差异性.这可能是由于指标对于不同等级的负荷的敏感度不同,即一些情绪性相关的指标可能对低负荷区域更敏感,而另一些情绪性相关的指标可能对高负荷区域更敏感.
- (3)利用(2)中所提到的阶段性显著差异指标与绩效-负荷倒U性关联建立线性分类模型,分类准确率超过93%,能够显著提升感知负荷的预测效果.

#### 参考文献

- [1] Huang K, Li J J, Wu Y M. A review of research on affective load theory and its application in the field of information behavior. *Libr Inf Serv*, 2018, 62(12): 116 (黄崑, 李京津, 吴英梅. 信息行为研究中的情感负荷理论及应用研究综述. 图书情报工作, 2018, 62(12): 116)
- [2] Sweller J. Cognitive load during problem solving: Effects on learning. *Cogn Sci*, 1988, 12(2): 257
- [3] Paas F G W C, Van Merriënboer J J G. Instructional control of cognitive load in the training of complex cognitive tasks. *Educ Psychol Rev*, 1994, 6(4): 351
- [4] Ginns P, Leppink J. Special issue on cognitive load theory: Editorial. *Educ Psychol Rev*, 2019, 31(2): 255
- [5] Sang L Q, Wang L, Qiao L, et al. Load-dependent activation of bilateral prefrontal cortex: An fNIRS study. J Army Med Univ, 2022(3): 210 (桑林琼, 王莉, 乔梁, 等. 基于 fNIRS 的大脑前额叶皮层在不同 脑力负荷下的激活模式研究. 陆军军医大学学报, 2022(3): 210)
- [6] Sweller J. Cognitive load theory [M] // Mestre J P, Ross B H. Psychology of learning and motivation. Pittsburgh: Academic Press, 2011: 37
- [7] Witt J A, Coras R, Becker A J, et al. When does conscious memory become dependent on the hippocampus? The role of memory load and the differential relevance of left hippocampal integrity for short- and long-term aspects of verbal memory performance. *Brain Struct Funct*, 2019, 224(4): 1599
- [8] Gilboa A, Marlatte H. Neurobiology of schemas and schemamediated memory. *Trends Cogn Sci*, 2017, 21(8): 618
- [9] Keller A S, Davidesco I, Tanner K D. Attention matters: How orchestrating attention may relate to classroom learning. CBE Life

- Sci Educ, 2020, 19(3): fe5
- [10] Nee D E. Integrative frontal-parietal dynamics supporting cognitive control. *eLife*, 2021, 10: e57244
- [11] Zheng Q Y. Cognition and emotion in multimedia learning—A perspective based on cognitive load theory. *Adv Soc Sci*, 2023(12): 6977
  (郑千屿. 多媒体学习中的认知与情感——基于认知负荷理论的研究视角. 社会科学前沿, 2023(12): 6977)
- [12] Plass J L, Kalyuga S. Four ways of considering emotion in cognitive load theory. Edu Psy Rev, 2019, 31: 339
- [13] Jia Z G. Research on the effect of mental workload on learning efficiency. *J Shaanxi Norm Univ* (*Philos Soc Sci Ed*), 2019, 48(4): 166 (贾智刚. 心理负荷对学习效率的影响研究. 陕西师范大学学报(哲学社会科学版), 2019, 48(4): 166)
- [14] Guo X D, Zheng H, Ruan D, et al. Associations between empathy and negative affect: Effect of emotion regulation. *Acta Psychol Sin*, 2023, 55(6): 892
  (郭晓栋, 郑泓, 阮盾, 等. 认知和情感共情与负性情绪: 情绪调节的作用机制. 心理学报, 2023, 55(6): 892)
- [15] Liu H S, Wang F, Zhang D. Inspiring Real-Time Evaluation and Optimization of Human–Robot Interaction with Psychological Findings from Human–Human Interaction. *Appl Sci*, 2023, 13(2): 676
- [16] Debie E, Fernandez Rojas R, Fidock J, et al. Multimodal fusion for objective assessment of cognitive workload: A review. *IEEE Trans Cybern*, 2021, 51(3): 1542
- [17] Wang Y Y, Zheng Y H. Multimodal data fusion: The core driving force to solve the key problems of intelligent education. *Mod Distance Educ Res*, 2022, 34(2): 93
  (王一岩, 郑永和. 多模态数据融合: 破解智能教育关键问题的核心驱动力. 现代远程教育研究, 2022, 34(2): 93)
- [18] Zheng L, Qiao X Q, Ni T, et al. Driver cognitive loads based on multi-dimensional information feature analysis. *China J Highw Transp*, 2021, 34(4): 240 (郑玲, 乔旭强, 倪涛, 等. 基于多维信息特征分析的驾驶人认知负荷研究. 中国公路学报, 2021, 34(4): 240)
- [19] Lorenz S, Helmert J R, Anders R, et al. UUX evaluation of a digitally advanced human–machine interface for excavators. *Multimodal Technol. Interact*, 2020, 4(3): 57
- [20] Liu R, Gao L L, Liu Y J, et al. Research on the usability evaluation model of human-computer interaction for command and control system // 2021 2nd International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI) .

  Shenyang, 2021: 213
- [21] Trisitichoke T, Ando S, Saito D, et al. Analysis of native listeners' facial microexpressions while shadowing non-native speech—
  Potential of shadowers' facial expressions for comprehensibility

- prediction // Interspeech 2019. Graz, 2019: 1861
- [22] Joseph A W, Murugesh R. Potential eye tracking metrics and indicators to measure cognitive load in human-computer interaction research. *J Sci Res*, 2020, 64(1):168
- [23] Xue Y F, Li Z W. Research on online learning cognitive load quantitative model based on eye tracking technology. *Mod Educ Technol*, 2019, 29(7): 59
  (薛耀锋, 李卓玮. 基于眼动追踪技术的在线学习认知负荷量化模型研究. 现代教育技术, 2019, 29(7): 59)
- [24] Wang X M. Three issues in cognitive load measurement: Constructs, observed indicators, and interpretations. *Glob Educ*, 2021, 50(6): 50
  (王小明. 认知负荷测量中的三大问题: 构念、观察指标与解释. 全球教育展望, 2021, 50(6): 50)
- [25] Li G F, Wu X J, Eichberger A, et al. Drivers' EEG responses to different distraction tasks. *Automot Innov*, 2023, 6(1): 20
- [26] Xue Y F, Wang K, Qiu Y S, et al. Cognitive load assessment of online learning based on multimodal data. *Mod Educ Technol*, 2024, 34(3): 79
  (薛耀锋, 王坤, 邱奕盛, 等. 基于多模态数据的在线学习认知负荷评估. 现代教育技术, 2024, 34(3): 79)
- [27] Chen B, Lu T Y, Yu L, et al. An accurate online testing model based on multimodal data. *Mod Educ Technol*, 2023, 33(4): 92 (陈波, 陆天易, 于冷, 等. 基于多模态数据的精准在线测试模型. 现代教育技术, 2023, 33(4): 92)
- [28] Xing R, Sun M R. The effects of negative emotion of different motivation types and cognitive load on working memory // Abstract Collection of the 25th National Conference on Psychology Group Presentation Reports. Chengdu, 2023: 704 (邢蓉, 孙美荣. 不同动机类型负性情绪和认知负荷对工作记忆的影响//第二十五届全国心理学学术会议摘要集——分组展贴报告. 成都, 2023: 704)
- [29] Xu S X, Wang Z L, Zhu S N, et al. Does cognitive load suppression based on working memory inhibit or enhance distraction interference. *Chin J Appl Psy*. 2024, 30(4): 298 (许双星, 王紫乐, 祝松楠, 等. 基于工作记忆的认知负荷抑制还是增强分心干扰. 应用心理学, 2024, 30(4): 298)
- [30] Zhao G J. The Influence of Negative Emotions on Sustained Attention [Dissertation]. Tianjin: Tianjin Normal University, 2023 (赵格静. 消极情绪对持续性注意的影响[学位论文]. 天津: 天津师范大学, 2023)
- [31] Pergher V, Vanbilsen N, Tournoy J, et al. Impact of strategy use during N-Back training in older adults. *J Cogn Psychol*, 2020, 32(8): 715
- [32] Makris A, Kosmopoulos D, Perantonis S, et al. Hierarchical feature fusion for visual tracking // 2007 IEEE International Conference on Image Processing. San Antonio, 2007: VI