

大学生图文信息知识整合的加工:来自 ERP 的证据*

赵笑梅 赵昕睿 王珂珂

(河北师范大学教育学院,石家庄 050024)

摘要:本研究采用改编的图形演示范式,通过行为实验和脑电技术考察两个独立但相关的图文信息知识整合的表现和加工过程,进一步明确了理论基础。结果表明:(1)无论是正确率还是反应时,图文与文图信息知识整合的表现均差于文文与图图信息知识整合;文本测验的整合表现优于图片测验;(2)图文信息知识整合发生在 300~500ms 的时间窗口内,不同呈现形式诱发的 N400 波幅存在显著差异,文文信息知识整合诱发的 N400 最大,图文与文图次之,图图最小。上述研究结果表明,信息的呈现形式影响知识整合,添加图片不利于图文信息知识整合。本研究明确了图文信息知识整合的加工过程,为图文信息知识整合支持多媒体学习的认知理论的观点提供了新的实验证据。

关键词:知识整合;图文信息;多媒体学习;大学生

分类号:B844

1 引言

知识整合是指将两个或多个独立但相关的信息联系起来自我衍生新知识的过程,它是积累知识库的有效模型(Bauer et al., 2020)。例如,在某个情节学到一个知识:“世界上最大的生态群落是北方森林”,在另一个情节学到一个与之相关的知识:“北方森林也叫做针叶林”,这两个知识可以自我衍生成一个新知识:“世界上最大的生态群落是针叶林”。其中,情节中学到的知识是主干事实,自我衍生的新知识是整合事实。研究表明,主干事实可以迅速整合(Bauer & Jackson, 2015),在大脑中长期保存(Varga & Bauer, 2013)。此外,知识整合还与学业成绩有关(Esposito & Bauer, 2017; Varga et al., 2019)。纵观以往研究,知识整合较多采用文本材料(Dugan & Bauer, 2022; Esposito & Bauer, 2019; 赵笑梅等, 2021),随着多媒体的普及,图文材料普遍应用多媒体教学中,两个独立但相关的图文信息整合表现及加工过程又如何呢?

图文整合是否促进学习值得深入探究。有关同一信息图文整合的研究显示,人们使用图文结合的材料学习时效果更好(Mayer, 2009; Mayer & Gallini, 1990),也有部分研究报告添加图片不利于获取知识(Magnus et al., 2020; Zhao et al., 2023)。而图文信息知识整合需联结两个独立的信息,与同一

信息的图文整合表现应有所不同。在图文信息知识整合领域,虽有研究者使用图片作为实验材料进行考察,但这些研究侧重完善整合阶段模型(Dugan et al., 2023)和比较不同年龄的发展差异(Miller-goldwater et al., 2021),并没有明确考察图文信息知识整合的表现及加工过程。在为数不多的图文知识整合研究中,Esposito 等人(2021)明确对比了文图与文文信息知识整合,结果发现儿童的整合表现无显著差异。然而该结果可能没有反映图文信息知识整合的真实表现。因为他们在呈现图片时,还提供了图片中重要词汇的音频和标签,这促进了理解和识记图文信息。图文信息知识整合的真实表现究竟如何是本研究的兴趣点之一。根据以往研究发现,知识整合在高表面相似性条件下更容易取得较好表现(Bauer et al., 2012; Bauer & Larkina, 2017; 赵笑梅等, 2025)。从表面相似性的角度看,相比较文文信息,文图信息是两种不同的呈现形式,其知识整合表现应该比文文更差。因此,本研究推测图文或文图信息知识整合表现比文文和图图信息知识整合差。

此外,图文信息知识整合是否也具有和同一信息图文整合相同或类似的理论基础也是值得探讨的问题。当前研究者主要用多媒体学习的认知理论和图文理解整合模型来解释同一信息的图文加工(Li et al., 2020; Magnus et al., 2020; Schüler et al.,

* 基金项目:河北省社会科学基金项目(HB20JY006)。

通讯作者:赵笑梅, E-mail: zhaoxiaomei@139.com

2017)。Mayer(2001)基于双重编码假设、有限容量假设和主动加工假设提出了多媒体学习的认知理论,并认为多媒体学习包括三个步骤:选择、组织和整合。首先,学习者需要选择图文的重要信息进入工作记忆中分别构建图像和声音;然后,工作记忆中的图像和声音组织成单一连贯的图像模型和言语模型;最后,学习者需将组织好的图像模型和言语模型,以及自己的先前知识经验整合成一个整体。Schnotz 和 Bannert(2003)对该理论提出了质疑,他们认为图文基于不同符号系统加工,且表征原则也有差异,不能进行简单的结构整合,于是提出了图文理解整合模型。该模型认为图文通过两种方式进行表征,一种是描绘性表征,即视觉表象和心理模型;另一种是描述性表征,即文字表征和命题表征。

两个理论的不同之处主要体现在加工过程和整合发生阶段。在加工过程上,Mayer(2009)假设图文在不同的认知子系统加工,两种加工建构两种平行的心理模型,最后彼此进行一对一映射联系;而 Schnotz 和 Bannert(2003)认为命题表征和心理模型之间不存在一对一映射关系,心理模型与命题表征存在持续互补关系。在整合发生阶段上,Mayer(2009)的理论强调言语模型与视觉模型以及长时记忆中先前知识的整合,整合发生在最后一步;但根据 Schnotz 和 Bannert(2003)的理论,第一个整合阶段应该发生在交互构建命题表征和心理模型时,第二个整合阶段发生在二者的交互整合中,这意味着整合可能早在图文进行深层加工时就发生了,而非等到图文独立加工后再整合。

目前两个理论的研究主要聚焦在同一信息图文整合上,但结论并不统一。Schnotz 等人(2014)用图文理解整合模型分析了中学生在整合图文时使用的策略,结果发现,图文构建心理模型的过程似乎是动态的,最初心理模型的构建主要由文本驱动,图片具有一定支架作用,该结论得到了眼动数据的支持。然而当前更多研究支持多媒体学习的认知理论。例如,Li 等人(2020)的研究表明,图文首先通过不同的神经通路加工,然后整合到共同的语义系统中,但该研究中被试的任务是判断句子与图片是否存在语义相关,并非知识整合。Schüler(2017)的研究也支持了多媒体学习的认知理论,结果显示,被试首先构建了两个独立的心理模型,然后将它们整合在一起,然而这种模式只在文本加工中观察到,在图片加工中没有观察到。如上述,虽然两个理论都有实证研究支持,但目前

尚缺乏直接检验图文信息知识整合的研究,哪个理论又能提供合理解释呢?

根据上述分析,两个理论的区别主要体现在加工过程和整合发生阶段。因此,本研究一方面通过行为实验得到图文信息知识整合表现,间接推测可能的加工过程;另一方面结合脑电技术探讨加工过程,获得整合发生的直接证据。

首先,加工过程的差异可以通过测验形式考察。本研究在图文和文图条件下使用文本测验和图片测验检验整合再认的成绩。按照多媒体学习的认知理论,个体在加工图文时建构了两种平行的心理模型,并建立了一一对应关系(Mayer, 1997),测验问题与文本选项的加工属于同一认知系统,提取时更容易,但其与图片选项的加工系统不同,将问题与图片表征联系起来具有一定难度,这可能导致图片测验结果差于文本测验。而按照图文理解整合模型的观点,命题表征和心理模型是图文信息交互形成的,那么这个持续交互的整体应不受测验形式的影响(Schnotz & Bannert, 2003)。

其次,整合阶段的差异可能体现在加工过程上。这需要借助事件相关电位(event-related potential, ERP)技术获得更为直接的证据,通过考察刺激识别到整合的整个认知过程(Li et al., 2020),可以明确其加工过程。根据多媒体学习的认知理论(Mayer, 2009)和图文理解整合模型(Schnotz & Bannert, 2003)以及神经影像学研究的成果(Li et al., 2020),可以推测图文早期都是在不同的通道获取和加工,在晚期进行整合(李松清等,2015)。然而该结论依据同一信息图文整合得出,图文信息知识整合的加工过程如何尚不清楚。鉴于知识整合的加工机制和以往研究(Li et al., 2020; 赵笑梅等, 2021),与语义加工密切相关的 N400 可能反映图文信息的整合过程(Kutas & Federmeier, 2011; 王穗莘,黄健,2019)。如果按照多媒体学习的认知理论的观点,整合发生在图文独立语义加工后,文图信息知识整合的 N400 可能更明显,其次是图文和文图,最后是图图。如果按照图文理解整合模型的观点,图文构建命题表征和心理模型时已完成了第一阶段的整合,不同呈现形式的 N400 应该不存在显著差异。除 N400 之外,N400 前后的脑电成分也可能揭示整合过程。在知识整合的研究中,P200 成分和 P600 可能反映了主干事实的语义分析(Bauer & Jackson, 2015; 赵笑梅等,2021)。基于以上叙述,本研究选择 150 ~ 300ms, 300 ~ 500ms 和 500 ~

800ms 时间窗进行观测。此外,成年人在编码时就能自动整合知识 (Bauer et al., 2021),本研究将重点考察呈现第二个主干事实后的 N400 与哪个理论的推测相符。

综上,虽有文献报告了图文整合的表现及认知加工,但图文信息知识整合表现如何尚不明确,其加工过程尚未得到深入探讨,两个理论在测验形式和脑电成分上的推测能否得到验证也需关注。为探究上述问题,本研究采用改编的图形演示顺序呈现图片或文本 (Esposito et al., 2021),并增加了文与图图条件作为对照组。本研究假设:(1)在行为结果上,图文与文图信息知识整合应该比文文与图图信息知识整合表现差;(2)如果图文信息知识整合支持多媒体学习的认知理论,则各条件在测验形式和脑电成分上存在显著差异;如果支持图文理解整合模型,则各条件在测验形式和脑电成分上不存在显著差异。

2 方法

2.1 被试

参考前人研究 (Li et al., 2020),采用 G * Pow-


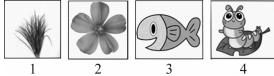

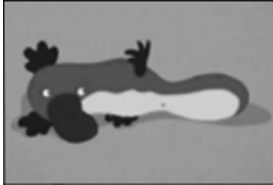
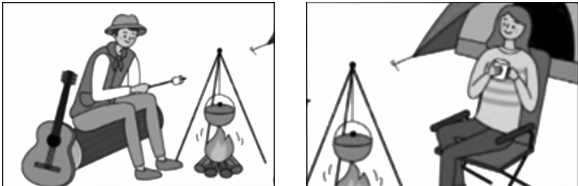
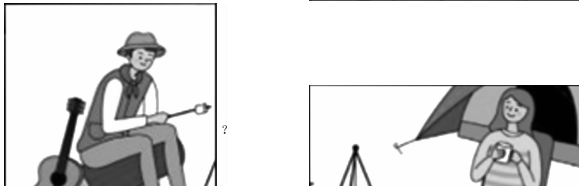


er 软件估算样本量,设置 $effect\ size\ f = 0.40, \alpha = 0.01, power = 0.95$,得到最小样本量为 20。本研究选择 30 名在校大学生 (女 17 名,男 13 名)参加实验,年龄 18 ~ 25 岁,平均年龄 21.61 ($SD = 2.15$) 岁。所有被试右利手,身体健康,实验前无身体不适,视力或矫正视力正常,色觉正常。所有被试自愿参加实验,并在实验结束后可以获得一定的报酬。

2.2 实验材料

实验材料内容和数量参照了 Schüler 等人 (2015) 和 Miller - goldwater 等人 (2021) 的研究,如表 1 所示。为保证学习材料是彼此相关的新颖知识,随机选取 16 名不参与正式实验的大学生对学习材料的匹配度 (1 = 非常不匹配,5 = 非常匹配) 和熟悉度 (1 = 非常熟悉,5 = 非常不熟悉) 进行 5 点量表评分,选取了评分在均值 4 分以上的实验材料。

学习阶段共有 68 个句子和 68 张图片。句子长度为 6 ~ 18 个字,图片为带有黑色边框的简单事物 (Scharinger et al., 2020), 每张图片大小为 250 × 170 像素。文文条件由 24 个 (12 对) 句子构成,每对可以整合成一个新的事实。例如,主干事实 1 为“哈利法塔是世界上最高的建筑”,主干事实 2 为“世

表 1 学习阶段与测验阶段的实验材料

呈现形式	主干事实 1	主干事实 2	整合再认测验	主干再认测验
文文	哈利法塔是世界上最高的建筑	世界上最高的建筑位于迪拜	哈利法塔位于? 1. 阿曼 2. 迪拜 3. 沙迦 4. 阿治曼	哈利法塔是世界上最高的建筑
图文		蜂鸟是唯一会倒着飞的鸟	会倒着飞的鸟的食物是? 	蜂鸟是唯一会倒着飞的鸟
文图	鸭嘴兽是体型最小的单孔目动物		最小的单孔目动物的嘴? 1. 扁平 2. 翘起来 3. 管型 4. 短粗	
图图				

界上最高的建筑位于迪拜”,二者可以整合成一个新知识:哈利法塔位于迪拜。图文条件由 12 个图片和 12 个句子(12 对)构成。例如,“蜂鸟吸食花蜜”的图片与“蜂鸟是唯一会倒着飞的鸟”的句子可以整合成一个新知识:会倒着飞的鸟的食物是花蜜。文图条件也是由 12 个句子和 12 个图片构成,学习过程与图文条件相同。图图条件由 24 张(12 对)图片构成,两张图片可以整合成一张新图片。四种条件的主干事实彼此不同。此外,每种条件分别有 20 个填充材料随机插入学习材料中,其作用在于干扰被试(Miller - goldwater et al., 2021),并不参与后续统计分析。

测验阶段包含整合再认测验和主干再认测验。整合再认测验中,文图和图图条件采用 12 个文本或 12 个图片测验,图文和文图条件均采用 6 个文本和 6 个图片测验。文本测验和图片测验均以句子提问,不同在于文本测验的选项为词语,图片测验的选项为图片,两种测验进行了顺序平衡。主干再认测验中,每种条件分别有 12 个文本或图片再认问题,被试需要判断文本或图片是否学过。

2.3 实验设计

本研究采用单因素被试内设计。自变量是信息呈现形式,共有四个水平:(1)文文;(2)图文;(3)文图;(4)图图。因变量是测验的正确率和反应时、ERP 平均波幅。

2.4 实验程序

实验程序分为学习和测验阶段。实验流程如图 1 所示。被试坐在安静的实验室中,佩戴 64 导电极帽,视线正对屏幕中央,距离显示器前约 90 厘米处进行实验。

学习阶段:正式实验前被试需要完成 4 个练习问题以保证充分理解指导语。正式学习时,首先向被试呈现 500ms 的注视点“+”。之后呈现主干事实 1,被试学习完该句子或图片后按空格键学习下一个主干事实 1,每学习完一个主干事实后呈现 1500ms 的空屏。学习完主干事实 1 后,进行 10 分钟的缓冲活动(Miller - goldwater et al., 2021),再呈现主干事实 2 和填充材料。为排除呈现先后顺序对结果的干扰,四种条件的顺序在被试间平衡,且每个整合条件中两组主干事实在被试间平衡。

测验阶段:每学完一个条件的主干事实后,在屏幕上呈现整合再认问题。同时会在下方呈现四个选

项,要求被试选择,随后向被试呈现主干事实,要求被试判断是否见过。被试作答不限时间。

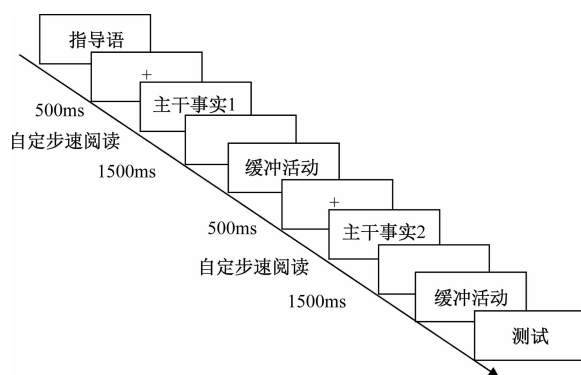


图 1 实验流程示意图

2.5 数据记录及分析

使用 NeuroScan 4.5 系统采集脑电信号。按照国际 10 - 20 系统扩展的 64 导电极帽记录数据,记录时以左侧乳突 M1 作为在线参考,前额接地。左眼上下电极记录垂直眼电,水平眼电由左右眼外侧电极记录,所有电极的头皮电阻均降至 5k Ω 以下,采样率为 500Hz/导。完成采集后离线处理数据,将所有 ERP 信号电压超过 $\pm 100\mu\text{V}$ 视作伪迹予以剔除。

用 MATLAB 2019b 的 EEGLAB V2021.1 和 ERPLAB 8.30 对数据进行预处理。首先,以双侧乳突的平均值做离线参考,带通滤波为 0.1 - 30Hz。其次,对刺激呈现前 150ms 和呈现后 800ms 进行分段,以刺激呈现前 150ms 作为基线进行校正。最后,用独立成分分析(ICA)的方法去除眼电和伪迹。在每一个电极上,把相同条件下所有被试的 ERP 信号进行叠加平均以便于进行后续的分析。

在 150 ~ 300ms、300 ~ 500ms、500 ~ 800ms 的时间窗内,计算每个条件下的 ERP 平均波幅,之后进行统计分析。分析因素有三个:呈现形式 \times 大脑半球 \times 脑区,大脑半球分别是右半球、中间半球和左半球,脑区分别是 F 区、FC 区、C 区、CP 区和 P 区。根据大脑半球和脑区将头皮分为 15 个区域:F 右(F2, F4, F6)、F 中(Fz)、F 左(F1, F3, F5);FC 右(FC2, FC4, FC6)、FC 中(FCz)、FC 左(FC1, FC3, FC5);C 右(C2, C4, C6)、C 中(Cz)、C 左(C1, C3, C5);CP 右(CP2, CP4, CP6)、CP 中(CPz)、CP 左(CP1, CP3, CP5);P 右(P2, P4, P6)、P 中(Pz)和 P 左(P1, P3, P5)。在方差分析中,球形假设结果显著时,采用 Greenhouse - Geisser 校正。

表 2 不同呈现形式条件下的知识整合表现和主干事实再认表现 ($M \pm SD$)

	整合再认表现		主干事实再认表现	
	正确率	反应时 (ms)	正确率	反应时 (ms)
文文	0.68 ± 0.15	2705.16 ± 420.18	0.88 ± 0.12	2301.61 ± 170.18
文图	0.55 ± 0.11	2800.52 ± 349.00	0.73 ± 0.09	2209.03 ± 343.34
图文	0.50 ± 0.11	3072.61 ± 420.28	0.69 ± 0.09	2385.82 ± 343.33
图图	0.86 ± 0.12	2729.95 ± 671.45	0.88 ± 0.76	2017.81 ± 339.02

3 结果

3.1 行为结果

删除反应时大于 2 个标准差之外的 3 个异常值,以正确率和反应时为因变量(表 2),对整合知识和主干事实再认表现进行重复测量方差分析。

在正确率上,整合再认表现差异显著, $F(3, 78) = 57.39, p < 0.001, \eta_p^2 = 0.69$ 。事后检验发现,图图显著优于文文 ($p < 0.001$)、图文 ($p < 0.001$) 和文图 ($p < 0.001$);文文显著优于图文 ($p < 0.001$) 和文图 ($p = 0.001$);文图和图文不存在显著差异, $p > 0.05$ 。主干再认表现差异显著, $F(3, 78) = 39.25, p < 0.001, \eta_p^2 = 0.60$ 。事后检验发现,图图和文文最好,显著优于图文 ($p < 0.001$) 和文图 ($p < 0.001$);文图和图文不存在显著差异, $p > 0.05$;图图和文文不存在显著差异, $p > 0.05$ 。

为了检验不同条件的整合表现是否受主干事实学习差异的影响。以呈现形式为自变量,整合再认正确率为因变量,主干事实再认正确率为协变量,进行协方差分析。结果显示,不同条件对整合正确率影响显著, $F(3, 103) = 28.73, p < 0.001, \eta_p^2 = 0.46$ 。

在反应时上,整合再认表现差异显著, $F(3, 78) = 3.21, p < 0.05, \eta_p^2 = 0.11$ 。事后检验发现,图文显著高于文文 ($p < 0.01$)、图图 ($p < 0.05$) 与文图 ($p < 0.05$);其他条件均无显著差异, $p > 0.05$ 。主干再认表现差异显著, $F(3, 78) = 8.02, p < 0.001, \eta_p^2 = 0.24$ 。事后检验发现,图图显著低于文文 ($p < 0.001$)、图文 ($p = 0.001$) 与文图 ($p < 0.05$);文图显著低于图文, $p < 0.05$;其他条件均无显著差异, $p > 0.05$ 。

表 3 图文和文图整合条件下图片测验与文本测验的表现 ($M \pm SD$)

	测验表现	
	正确率	反应时 (ms)
图片测验	0.37 ± 0.12	3743.13 ± 889.13
文本测验	0.58 ± 0.21	2760.43 ± 469.03

最后,为了考察测验形式对整合结果的影响,对图文和文图条件下的文本测验和图片测验结果(表 3)进行差异检验。结果显示,文本测验正确率显著高于图片测验, $F(1, 26) = 20.06, p < 0.001, \eta_p^2 = 0.44$ 。文本测验反应时显著低于图片测验, $F(1, 26) = 37.14, p < 0.001, \eta_p^2 = 0.59$ 。

3.2 脑电结果

对不同呈现形式所诱发的脑电平均波幅进行描述统计,结果如表 4。波形图与地形图如图 2 与图 3 所示。

表 4 不同呈现形式诱发的 ERP 平均波幅 (μV)

呈现形式	150 ~ 300ms	300 ~ 500ms	500 ~ 800ms
文文	5.93 ± 2.72	5.50 ± 5.93	3.50 ± 0.92
文图	2.16 ± 1.73	5.11 ± 4.59	5.01 ± 0.41
图文	3.23 ± 3.58	4.40 ± 5.43	5.31 ± 2.86
图图	0.05 ± 0.70	1.08 ± 2.55	3.03 ± 1.13

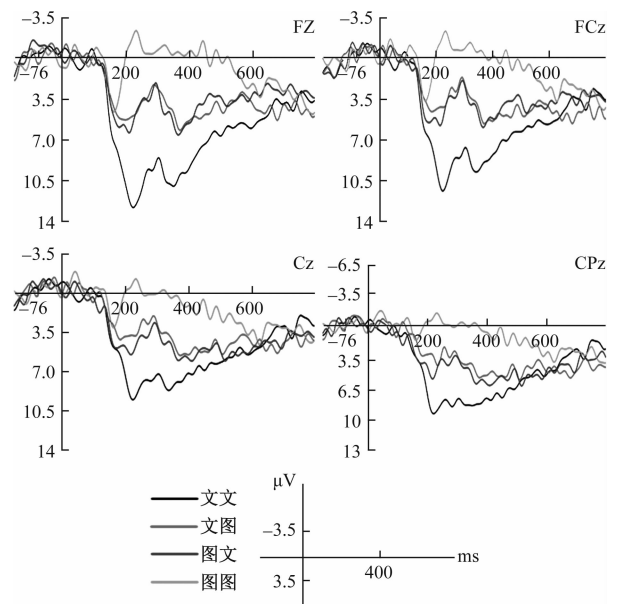


图 2 不同呈现形式在各脑区上的波形图

3.2.1 刺激呈现 150 ~ 300ms 时间窗口

对呈现形式 × 大脑半球 × 脑区进行重复测量方差分析,结果显示,半球主效应显著, $F(2, 54) =$

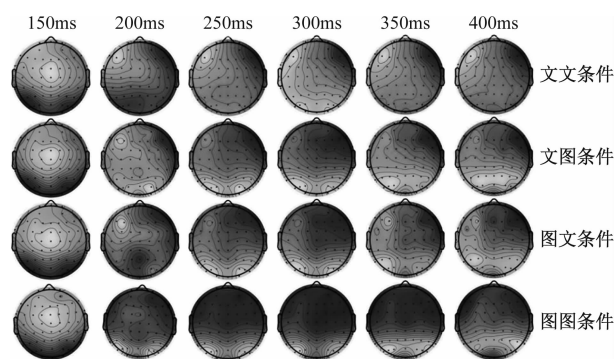


图3 不同信息呈现形式的地形图

21.43, $p < 0.05$, $\eta_p^2 = 0.44$; 呈现形式主效应不显著, $F(2.01, 54.28) = 7.85$, $p > 0.05$, $\eta_p^2 = 0.26$; 脑区主效应不显著, $F(1.33, 35.94) = 3.842$, $p > 0.05$, $\eta_p^2 = 0.16$ 。呈现形式与脑区交互作用显著, $F(2.55, 68.91) = 25.85$, $p = 0.01$, $\eta_p^2 = 0.49$; 脑区和半球交互作用显著, $F(2.26, 60.98) = 22.05$, $p < 0.05$, $\eta_p^2 = 0.45$; 呈现形式与半球交互作用不显著, $F(2.55, 68.91) = 25.85$, $p > 0.05$, $\eta_p^2 = 0.26$ 。呈现形式、脑区和半球三阶交互作用不显著, $F(2.21, 59.71) = 15.92$, $p > 0.05$, $\eta_p^2 = 0.37$ 。简单效应分析显示, 不同呈现形式在 F 区 ($p < 0.05$)、FC 区 ($p = 0.05$) 的平均波幅有显著差异。在两个脑区中, 图图的波幅最小, 其次是图文和文图, 文文的波幅最大, 在其他脑区上的波幅无差异; 而各个脑区的平均波幅在文文 ($p = 0.01$)、图文 ($p < 0.01$) 和文图 ($p < 0.05$) 上有显著差异, 在图图上无显著差异。这些结果表明, 图文信息的初步语义加工在额叶区引起了不同的神经活动, 早期图文信息加工相对独立。

3.2.2 刺激呈现 300 ~ 500ms 时间窗口

结果显示, 脑区主效应显著, $F(1.29, 34.85) = 57.94$, $p < 0.01$, $\eta_p^2 = 0.68$; 呈现形式主效应不显著, $F(1.69, 45.71) = 7.02$, $p > 0.05$, $\eta_p^2 = 0.10$; 半球主效应不显著, $F(1.37, 37.01) = 14.16$, $p > 0.05$, $\eta_p^2 = 0.34$ 。呈现形式与脑区交互作用显著, $F(2.24, 60.45) = 17.79$, $p < 0.05$, $\eta_p^2 = 0.40$; 呈现形式与半球交互作用不显著, $F(2.35, 63.34) = 8.02$, $p > 0.05$, $\eta_p^2 = 0.23$; 脑区和半球交互作用不显著, $F(1.86, 50.19) = 15.89$, $p > 0.05$, $\eta_p^2 = 0.37$ 。呈现形式、脑区和半球三阶交互作用不显著, $F(4.56, 191.69) = 2.31$, $p > 0.05$, $\eta_p^2 = 0.26$ 。简单效应分析显示, 不同呈现形式在 F 区 ($p < 0.01$) 的平均波幅有显著差异。文文的 N400 更大, 图图比

图文和文图的 N400 更小, 在其他脑区上的波幅无差异。这些结果表明, 早期语义加工后, 图文语义信息加工在额叶区存在显著差异, 文本加工程度更深。

3.2.3 刺激呈现 500 ~ 800ms 时间窗口

结果显示, 脑区主效应不显著, $F(2.12, 57.31) = 14.22$, $p > 0.05$, $\eta_p^2 = 0.35$; 半球主效应不显著, $F(1.49, 40.26) = 1.66$, $p > 0.05$, $\eta_p^2 = 0.09$ 。呈现形式与脑区交互作用不显著, $F(1.73, 46.87) = 2.72$, $p > 0.05$, $\eta_p^2 = 0.09$; 呈现形式与半球交互作用不显著, $F(2.59, 69.87) = 1.63$, $p > 0.05$, $\eta_p^2 = 0.06$; 脑区和半球交互作用不显著, $F(2.58, 69.57) = 6.57$, $p > 0.05$, $\eta_p^2 = 0.20$ 。呈现形式、脑区和半球三阶交互作用不显著, $F(4.56, 191.69) = 2.31$, $p > 0.05$, $\eta_p^2 = 0.13$ 。这些结果表明, 此时间窗口前已完成了图文信息加工。

4 讨论

本研究考察了图文信息知识整合的表现和加工过程, 并探讨了理论基础。行为数据显示, 添加图片不利于图文信息知识整合; 整合表现受测验形式影响, 文本测验正确率显著高于图片测验。脑电数据显示, 在 150 ~ 300ms, 图文进行早期语义加工, 图片的语义加工相对较缓慢; 在 300 ~ 500ms, 知识整合完成, 但根据图片的语义信息整合知识有一定难度。

4.1 图文信息知识整合表现

行为实验发现, 图文与文图表现较差, 这与既往相关研究不同。本研究文文和图图表现显著好于图文和文图, 图图正确率最高, 其次是文文, 最后是图文和文图, 四种条件的反应时也显示出相似趋势。具体来说, 图图表现最好可能是由于其直观性最强, 更容易被感知和记住(胡成霞, 2023)。单一呈现形式的知识整合表现显著好于图文结合的形式, 这与 Esposito 等人(2021)的研究不一致, 他们发现儿童文图与文文表现没有显著差异。这可能是由于 Esposito 等人(2021)的实验材料并不是单纯呈现图或文, 图片中额外的音频和标签提供了提示, 促进了语义加工。而且本研究与该项研究最大的区别是采用了大学生被试而不是儿童。Mayer 和 Gallini(1990)发现, 在先验知识水平较低的情况下, 图形演示有助于任务表现, 这表明儿童可能比大学生更受益于图片。

根据前人研究, 图文信息知识整合表现较差的原因可以归结为两方面: 一方面, 图文是低表面相似性材料。当表面相似性低时, 个体很难识别图文之

间整合元素的联系,这对知识整合是不小的困难(Bauer et al., 2021; Esposito & Bauer, 2019)。另一方面,可能与图文信息特定的加工策略有关。在图文和文图条件下,个体需要从图文材料中寻找对应的整合元素建立联系并生成新知识。Menendez (2023)的研究表明,图片包含丰富信息,当整合元素不明确时,更容易出现注意资源竞争,这导致个体更倾向将注意力集中在文本上,忽略图片细节,对图文信息中的整合元素建立联系带来挑战。此外,本研究还发现图文与文图在反应时上存在显著差异。这可能是由于视觉编码信息比视觉表征信息保持的时间稍长(Coltheart, 1980),先呈现文本有利于后续知识整合。因此,图文呈现顺序加大了知识整合难度,但两种呈现顺序的正确率并没有显著差异。

由于知识整合建立在学习两个主干事实的基础上,图文与文图表现差是否由学习上的差异导致也需探讨。本研究发现,图文和文图的主干再认测验成绩显著差于文文和图图。但协方差分析显示,不同条件下整合知识的正确率不能解释为主干事实学习差异的影响,这可能与提取引起的遗忘现象有关(赵丽波等,2017)。由于被试在整合再认测验中需要从大量记忆信息中提取整合知识,使得随后主干再认测验中相关但非目标的主干事实变得难以提取。而图文和文图的整合难度较大,进一步加剧了整合知识与主干事实的竞争(杨邵峰等,2022),最终导致图文与文图的主干再认测验成绩较差。

在检验理论方面,本研究在一定程度上支持了多媒体学习的认知理论。行为结果发现,图文信息知识整合受测验形式影响。在图文和文图条件下,文本测验表现显著优于图片测验。可能是由于文本与图片测验都以句子提问,图文信息形成表征后才建立一一对应的关系(Mayer, 1997),测验问题与文本表征在同一系统加工,更容易提取。对于图片测验来说,图文作为独立表征进行语义整合,图片的表征和问题的加工系统不同,被试将问题与具体的图片心理表征联系起来的表现反而差于文本测验,这符合多媒体学习的认知理论的推测。按照图文理解整合模型的观点,当呈现图文时,图文深层加工时存在持续交互,Schüler等人(2015)认为这可以看作是构建了包含图文信息的单一心理表征,那么文本测验与图片测验的表现应无显著差异。

4.2 图文信息知识整合的加工过程

虽然不少研究都关注了同一信息图文整合的加工过程和理论,但图文信息知识整合加工过程和理

论基础的研究还存在不足。本研究除了明确加工过程外,还通过分析加工过程反过来为两个理论对图文信息知识整合的解释提供了较为直接的证据。

脑电结果揭示了图文信息加工过程。在150~300ms,图文信息加工独立进行(李松清等,2015),个体对图片加工形成图像表征,对文本加工形成言语表征,并将这两种表征暂存在工作记忆中(Mayer, 2001);在300~500ms,图文信息可能已经发生了整合;在500ms前,整合已经完成。

在200ms左右,四种条件的脑电波幅存在显著差异。文文最大,其次是文图,之后是图文,图图最小,其中文图和图文差异不大。根据以往研究认为该成分是P200(Stuellein et al., 2016),然而当前对于P200的意义存在不同观点,有研究认为P200与词汇的语义加工有关(王彭等,2022;张积家,王斌,2024;张钦等,2003),也有研究认为这一成分反应了早期对图片的选择性注意(Johnen & Harrison, 2020)。如果该成分反映了对图片的注意,那么图文和文图的波幅可能会高于文文,然而结果中并未观察到这一现象。与文本相比,图片具有更生动的意义通达,个体对图片进行语义加工相对较慢。因此,P200最有可能反映了图文的早期语义加工。

在300~500ms,呈现形式与脑区差异显著,文文的N400最大,其次是图文和文图,最后是图图。有研究表明N400是对即将出现的词语和句子之间映射困难程度的反应(Kutas & Federmeier, 2011),是对语义整合过程的反应(Hagoort et al., 1999)。个体在对图文信息进行加工时,图文对应关系难度大于文文,这可能是由于文文的整合元素比图文和文图更为明显。而图图的N400最小,可能是由于个体在看到两个相关的图片时,试图进行语义整合,但难度较大。此外,还需考虑N400的差异是否由图文材料在物理属性的差异所导致,而不是整合发生阶段的差异。Yum等人(2011)的研究发现,图文材料的物理属性对早期成分有影响,对晚期成分影响不大,N400的加工依赖于刺激意义。因此,图文信息知识整合在N400上的差异并不是由于刺激的物理属性所导致(Li et al., 2020)。

N400的差异与多媒体学习的认知理论的假设一致,这为支持该理论提供了直接证据。N400的差异可能反映了不同条件下语义整合的加工难度(Kutas & Federmeier, 2011)。由于不同信息在早期独立进行语义加工并产生了差异,直到整合发生前,信息都没有交互,因此不同条件的语义整合难度存

在差异,这符合多媒体学习的认知理论的假设。而按照图文理解整合模型的观点,个体在整合文本和图片时形成了命题表征和心理模型,持续交互下的语义加工程度应该保持相对一致,呈现形式间的 N400 应该不存在显著差异,然而研究并没有发现这一点。

在 500 ~ 800ms,主效应和交互作用均不显著。Brouwer 等人(2012)认为 P600 是对整合加工过程的反应,由于其在时间窗口上比 N400 更为延迟,更倾向于对语义异常的重新分析和检查语义冲突的体现(Brouwer et al., 2012; Kim et al., 2018)。在本研究中,图文信息并不存在歧义,并没有诱发 P600。此外,Brouwer 等人(2012)也指出,与 N400 有关的语义整合难度也可能反映在 P600 上,但该时间窗口并没有发现 P600,说明知识整合在 500ms 前就已完成。

4.3 研究意义、局限与展望

本研究从行为和脑电实验两个层面探究了图文信息知识整合表现及加工过程,对探讨其理论基础做了创新尝试,研究结果为多媒体学习的认知理论提供了行为和神经生理证据。证据体现在测验形式和 N400 的差异上。多媒体学习的认知理论指出,个体对图文分别加工形成相应表征,随后与先前知识经验整合在一起(Mayer, 2001)。本研究结果发现图文信息知识整合受测验形式影响,说明图文加工时相对独立,而不是在深层加工时相互整合,这一结果与 Li 等人(2020)的研究相一致。在加工过程方面,由于图文独立加工,整合时图文的语义加工程度也有所不同,说明图文深层加工时并没有持续互动。

本研究的不足在于采用了少部分黑白图片作为实验材料,当物体颜色与记忆中的不一致时,可能会诱发更大的 N400(周文洁等,2021);虽然本研究仅用了几张黑白图片,对研究结果没有产生较大影响,但未来研究仍需注意。另外,本研究只考察了两个新学习的知识,未来还可以进一步探讨新旧图文信息的知识整合(赵笑梅等,2021)。再有,本研究的 ERP 数据是基于一致图文信息的分析,当个体在编码不一致的图文信息时,大脑产生了怎样的神经振荡,或许可以发掘整合的具体时间和脑区,这是未来值得进一步分析的问题。

5 结论

本研究探讨了图文信息知识整合的表现及加工

过程,进一步探讨了理论基础。行为实验发现添加图片不利于图文信息知识整合。脑电实验发现图文信息知识整合发生在 300 ~ 500ms 的时间窗口内,不同呈现形式诱发的 N400 波幅存在显著差异,文图信息知识整合诱发的 N400 最大,图图与文图次之,图图最小。本研究支持了多媒体学习的认知理论的观点。

参考文献:

- Bauer, P. J., & Jackson, F. L. (2015). Semantic elaboration: ERPs reveal rapid transition from novel to known. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(1), 271–282.
- Bauer, P. J., & Larkina, M. (2017). Realizing relevance: The influence of domain-specific information on generation of new knowledge through integration in 4 to 8 year old children. *Child Development*, 88(1), 247–262.
- Bauer, P. J., Cronin-Golomb, L. M., Porter, B. M., Jaganjac, A., & Miller, H. E. (2021). Integration of memory content in adults and children: Developmental differences in task conditions and functional consequences. *Journal of Experimental Psychology: General*, 150(7), 1259–1278.
- Bauer, P. J., Esposito, A. G., & Daly, J. J. (2020). Self-derivation through memory integration: A model for accumulation of semantic knowledge. *Learning and Instruction*, 66(1), 101271. <https://doi.org/10.1016/j.learninstruc.2019.101271>
- Bauer, P. J., King, J. E., Larkina, M., Varga, N. L., & White, E. A. (2012). Characters and clues: Factors affecting children's extension of knowledge through integration of separate episodes. *Journal of Experimental Child Psychology*, 111(4), 681–694.
- Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about semantic illusions: Rethinking the functional role of the p600 in language comprehension. *Brain Research*, 1446(1), 127–143.
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27(3), 183–228.
- Dugan, J. A., & Bauer, P. J. (2022). Putting the pieces together: Cognitive correlates of self-derivation of new knowledge in elementary school classrooms. *Journal of Experimental Child Psychology*, 221, 105441. <https://doi.org/10.1016/j.jecp.2022.105441>
- Dugan, J. A., Lee, K., Hanft, M. H., & Bauer, P. J. (2023). Selection of information necessary for successful self-derivation. *Cognitive Development*, 68(2023), 101364. <https://doi.org/10.1016/j.cogdev.2023.101364>
- Esposito, A. G., & Bauer, P. J. (2017). Going beyond the lesson: Self-generating new factual knowledge in the classroom. *Journal of Experimental Child Psychology*, 153, 110–125.
- Esposito, A. G., & Bauer, P. J. (2019). From bench to classroom: Collaborating within a dual-language education model. *Journal of Cognition and Development*, 20(2), 165–181.
- Esposito, A. G., Lee, K., Dugan, J. A., Lauer, J. E., & Bauer, P. J. (2021). Relating a picture and 1000 words: Self-derivation

- through integration within and across presentation formats. *Cognitive Development*, 60, 101099. <https://doi.org/10.1016/j.cogdev.2021.101099>
- Hagoort, P., & Brown, C. M. (1999). Gender Electrified: ERP Evidence on the Syntactic Nature of Gender Processing. *Journal of Psycholinguistic Research*, 28(6), 715–728.
- Johnen, A. -K., & Harrison, N. R. (2020). Level of uncertainty about the affective nature of a pictorial stimulus influences anticipatory neural processes: An event-related potential (ERP) study. *Neuropsychologia*, 146(2020), 107525. <https://doi.org/10.1016/j.neuropsychologia.2020.107525>
- Kim, A. E., Oines, L., & Miyake, A. (2018). Individual differences in verbal working memory underlie a tradeoff between semantic and structural processing difficulty during language comprehension: An ERP investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(3), 406–420.
- Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*, 62(1), 621–647.
- Li, S., Chen, S., Zhang, H., Zhao, Q., Zhou, Z., Huang, F., ... Hong, J. (2020). Dynamic cognitive processes of text-picture integration revealed by event-related potentials. *Brain Research*, 1726, 146513. <https://doi.org/10.1016/j.brainres.2019.146513>
- Magnus, L., Schütte, K., & Schwanewedel, J. (2020). Challenges Solving Science Tasks with Text-Picture Combinations Persist beyond Secondary School. *Journal of Research on Educational Effectiveness*, 13(4), 759–783.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist*, 32(1), 1–19.
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: Cambridge University Press.
- Mayer, R. E., & Gallini, J. K. (1990). When Is an Illustration Worth Ten Thousand Words? *Journal of Educational Psychology*, 82(4), 715–726.
- Menendez, D. (2023). Cues to generality: Integrating linguistic and visual information when generalizing biological information. *Journal of Educational Psychology*, 115(8), 1110–1124.
- Miller-Goldwater, H. E., Cronin-Golomb, L. M., Porter, B. M., & Bauer, P. J. (2021). Developmental differences in reactivation underlying self-derivation of new knowledge through memory integration. *Cognitive Psychology*, 129, 101413. <https://doi.org/10.1016/j.cogpsych.2021.101413>
- Scharinger, C., Schüler, A., & Gerjets, P. (2020). Using eye-tracking and EEG to study the mental processing demands during learning of text-picture combinations. *International Journal of Psychophysiology*, 158, 201–214.
- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning & Instruction*, 13(2), 141–156.
- Schnotz, W., Ludewig, U., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2014). Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology*, 106(4), 974–989.
- Schüler, A. (2017). Investigating gaze behavior during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction*, 49, 218–231.
- Schüler, A., Arndt, J., & Scheiter, K. (2015). Processing multimedia material: Does integration of text and pictures result in a single or two interconnected mental representations? *Learning and Instruction*, 35, 62–72.
- Stuellein, N., Radach, R. R., Jacobs, A. M., & Hofmann, M. J. (2016). No one way ticket from orthography to semantics in recognition memory: N400 and P200 effects of associations. *Brain Research*, 1639, 88–98.
- Varga, N. L., Esposito, A. G., & Bauer, P. J. (2019). Cognitive correlates of memory integration across development: Explaining variability in an educationally relevant phenomenon. *Journal of Experimental Psychology: General*, 148(4), 739–762.
- Varga, N. L., & Bauer, P. J. (2013). Effects of delays on 6-year-old children's self-generation and retention of knowledge through integration. *Mind, Brain, and Education*, 115(2), 326–341.
- Yum, Y. N., Holcomb, P. J., & Grainger, J. (2011). Words and pictures: An electrophysiological investigation of domain specific processing in native Chinese and English speakers. *Neuropsychologia*, 49(7), 1910–1922.
- Zhao, F., Gashler, R., Wagner, I., & Schnotz, W. (2023). Lower grade students tend to give up early in multimedia learning. *European Journal of Psychology of Education*, 38(2), 545–565.
- 胡成霞. (2023). 从图文关系嬗变看教科书插图的价值定位及实现路径·课程·教材·教法, 43(11), 35–42.
- 李松清, 赵庆柏, 周冶金, 张依. (2015). 多媒体学习中图文加工的认知神经机制. *心理科学进展*, 23(8), 1361–1370.
- 王彭, 李培, 张珊珊, 杨亦鸣. (2022). 汉语空缺动词语义加工的 ERPs 证据——兼谈汉语动词空缺句. *外国语(上海外国语大学学报)*, 45(1), 14–24.
- 王穗苹, 黄健. (2019). 语言理解中的语义加工: 不同模态神经影像的研究. *生理学报*, 71(1), 127–139.
- 杨邵峰, 张志超, 冀婷. (2022). 抑制还是干扰? 来自提取练习范式的证据. *心理与行为研究*, 20(6), 753–759.
- 张积家, 王斌. (2024). 义符认知功能的心理语言学探索——三十年研究工作之回顾. *西北师大学报(社会科学版)*, 61(01), 104–117.
- 张钦, 丁锦红, 郭春彦, 王争艳. (2003). 名词与动词加工的 ERP 差异. *心理学报*, 35(6), 753–760.
- 赵丽波, 张悦, 于玮雯, 邓丽芳. (2017). 提取引起的遗忘的认知机制. *心理科学进展*, 25(10), 1738–1748.
- 赵笑梅, 靳伟琼, 王莹平, 褚铮. (2021). 知识整合的时间进程: ERP 研究. *心理科学*, 44(5), 1089–1096.
- 赵笑梅, 程释, 刘子涵, 刘红. (2025). 不同语言条件和执行功能对大学生记忆整合的影响. *心理发展与教育*, 41(2), 153–162.
- 周文洁, 邓丽群, 丁锦红. (2021). 物体颜色对情景记忆的影响. *心理学报*, 53(3), 229–243.

The Processing of Text-graphic Information Knowledge Integration of College Students: Evidence from ERP

ZHAO Xiaomei ZHAO Xinrui WANG Keke

(College of Education, Hebei Normal University, Shijiazhuang 050024)

Abstract: This study utilized an adapted graphic presentation paradigm to investigate the performance of two separate yet related knowledge integrations involving graphic and text information, and the theoretical basis was further clarified. The investigation involved conducting behavioral experiments and ERP technology. The findings reveal the following: (1) The form of information presentation influences knowledge integration. Whether it pertains to accuracy rate or response time, the performance of integrating graphic-text and text-graphic is inferior to that of text-text and graphic-graphic integration. Moreover, the integration performance of text-based tests surpasses that of graphic-based tests; (2) The knowledge integration of graphic and text information occurs within the time frame of 300 to 500ms. Notably, there exist significant differences in the N400 amplitudes induced by various presentation forms. The N400 induced by text-text integration is the largest, followed by graphic-text integration and text-graphic integration. The research results presented above demonstrate the impact of information presentation on knowledge integration. It is observed that the inclusion of graphics does not promote the integration of graphic and text information. This study offers valuable insights into the process of knowledge integration between graphic and text information, and provides experimental evidence that supports the cognitive theory of multimedia learning's perspective on the integration of graphic and text information.

Key words: knowledge integration; text-graphic information; multimedia learning; college students