心理学报 2026, Vol. 58, No. 12, 2520–2531 https://doi.org/10.3724/SP.J.1041.2026.2520
Acta Psychologica Sinica https://cstr.cn/32110.14.2026.2520

指令记忆中被试执行任务效应的多模态依赖：指令惯用性的调节作用*

谢婷婷[1] 杨联宇[1] 吴晓甜[1] 周典典[1] 于战宇[2]

([1]福建师范大学心理学院, 福州 350117) ([2]江苏师范大学教育学部, 徐州 221116)

摘 要 被试执行任务(Subject-Performed Task, SPT)效应指个体以动作编码(执行指令所描述的动作)而非言语编码(仅听指令)来学习指令时, 其后续回忆准确性更高。本研究从多通道视角解释 SPT 效应, 以三项实验考察指令惯用性对 SPT 效应多通道依赖性的调节作用。实验 1 操控指令惯用性与物体图片可见性, 发现惯用指令在有无图片条件均产生 SPT 效应, 而非惯用指令均未出现 SPT 效应。实验 2 将图片改为真实物体并允许操作, 发现非惯用指令在此条件出现 SPT 效应。实验 3 正交操纵视觉(戴眼罩)与触觉(戴手套), 发现非惯用指令的 SPT 效应仅在视觉可用时出现, 触觉的精细信息不影响 SPT 效应。三项实验表明, SPT 效应的通道依赖模式随指令惯用性而不同: 高惯用指令由动觉独立驱动, 低惯用指令需视觉–动觉协同, 触觉精细信息则非必需。

关键词 SPT 效应, 指令记忆, 多通道加工理论

分类号 B842

1 引言

指令记忆(memory for verbal instructions)指个体对言语指令(如“推勺子”)进行即时编码、保持与提取的能力(Yang et al., 2018)。作为课堂听讲、任务执行与生活自理的基础认知能力(Wang et al., 2022; Xie et al., 2024), 其水平直接影响个体的适应性行为表现。研究表明动作编码可以提高指令记忆: 当学习者以动作编码(即亲手执行指令所描述的动作)而非言语编码(即仅听指令)来学习一系列言语指令时, 其后续回忆准确性更高; 这一现象被称为被试执行任务效应(Subject-Performed Task effect, SPT 效应; Allen et al., 2019; Cohen, 1981)。

关于 SPT 效应的产生机制, 多通道加工理论(multimodal theory)提出, 动作编码优于言语编码, 是因为其在言语通道之外, 额外激活动觉、视觉(如观察物体与动作)、触觉(如接触物体表面)等多元感觉通道, 从而增强记忆痕迹的丰富性与可提取性(王丽娟, 李广政, 2014; Bäckman & Nilsson, 1984; Engelkamp & Zimmer, 1984)。然而, 该理论尚未阐明各通道的相对贡献。在动作编码可能激活的多个感觉通道中, 视觉(提供物体的视觉信息)和触觉(提供物体的质地、重量等接触信息)是与物体交互时最直接、最常被激活的两个通道, 厘清二者的作用是理解 SPT 效应多通道机制的关键切入点。

已有研究通过操控物体可见性来检验视觉通道的作用, 但结果存在分歧: 部分研究发现 SPT 效应在无实物条件下存在(Engelkamp, 1996; Kormi-Nouri, 2000; Nyberg et al., 1991), 另一些则报告该效应仅在物体可见时出现(Mecklenbräuker et al., 2011)。对于此不一致, 于战宇(2014)提出材料熟悉度可能是调节变量。该研究操纵指令中名词的视觉熟悉度(即被试是否知晓该名词所指物体的样子), 发现对于低熟悉名词组成的言语指令, SPT 效应仅在呈现物体图片时出现; 而对高熟悉名词组成的言语指令, SPT 效应不受物体图片可见性影响。

收稿日期: 2026-01-26

* 教育部人文社科基金(24YJC190035)和福建省自然科学基金(2026J008125)资助。

通信作者: 于战宇, E-mail: yzhany@jsnu.edu.cn

值得注意的是, 于战宇(2014)研究中所有言语指令均为惯用动宾搭配(如高熟悉度的“拉拉链”、低熟悉度的“调琴弦”), 动词与名词已形成高度自动化的联结。当物体名词为个体所熟悉时, 对由其构成的惯用动宾搭配的言语指令(以下简称“惯用指令”)的记忆可能依赖内部激活的多通道表征(包括运动程序、典型视觉场景与触觉预期), 而较少依赖外部物体; 若个体对物体名词本身不熟悉, 则需借助物体来唤起相应表征。近期研究显示, 即使在无实物条件下, 对由熟悉物体名词构成的惯用指令仍可观测到显著 SPT 效应(王丽娟, 马嘉琳, 2023; Ma et al., 2025), 一定程度上支持“对于由熟悉物体名词构成的惯用指令, 内部线索足以支撑 SPT 效应”的推测。然而, 当指令为非惯用搭配(如“转衣架”)(以下简称“非惯用指令”)时, 情况可能不同。这类指令的动词与名词缺乏既有联结, 需即时整合动词与名词, 其记忆过程可能更依赖外部感觉输入。换言之, 物体是否呈现可能决定这类指令下 SPT 效应的出现。尽管有研究曾使用非惯用指令作为实验材料(如 Nyberg et al., 1991), 但其实验未操控物体可见性, 因此一个尚未解决的问题是: 对于由熟悉物体名词构成的指令, 当动词与名词为非惯用搭配时, 外部视觉输入是否是 SPT 效应产生的必要条件。

此外, 现有研究在考察“物体呈现”对 SPT 效应的影响时, 普遍将物体作用归因于视觉线索。于战宇(2014)使用物体图片, 主要提供视觉线索。然而, 多数其他研究(Engelkamp, 1996; Kormi-Nouri, 2000; Mecklenbräuker et al., 2011; Nyberg et al., 1991)使用真实物体, 同时提供视觉与触觉输入。这些研究将“物体呈现”的效应归因于视觉通道, 可能忽略了触觉信息(如物体的质地、重量、形状)的潜在贡献。触觉同样可能通过增强动作表征的丰富性, 对 SPT 效应产生作用。鉴于多通道加工理论明确将触觉列为潜在贡献通道(Bäckman & Nilsson, 1984), 而其独立作用尚未被分离, 有必要专门考察触觉输入对 SPT 效应的特异性贡献, 以更精确地验证多通道理论的预测。

本研究旨在从多通道加工理论框架考察 SPT 效应的产生机制, 拟回答两个问题: (1)对于由熟悉物体名词(如“铅笔”)构成的指令, 惯用搭配(如“削铅笔”)与非惯用搭配(如“戳铅笔”)在动作编码下诱发的 SPT 效应是否表现出不同的感觉通道依赖性? (2)触觉输入是否对 SPT 效应具有独立贡献? 厘清这两个问题, 有助于澄清多通道加工理论中“多通道协同”的具体方式, 即是各通道都有作用且通道越多作用越大, 还是说通道的作用因指令能否唤起已有动作经验而不同, 亦可为视觉或触觉功能受限人群(如孤独症儿童、视障/触觉损伤个体)提供更具针对性的记忆支持策略。

为了回答这两个问题, 本研究设计了三个实验。实验 1 采用生活中常见的熟悉物体, 通过动词操控指令的惯用性, 并如同于战宇(2014)的研究, 通过物体图片的有无操纵视觉通道, 检验: 在 SPT 效应对视觉通道的依赖上, 惯用指令与非惯用指令是否不同。由于惯用指令的动词与名词之间的既有联结很强, 我们假设其 SPT 效应的产生不依赖视觉通道, 仅凭动觉输入即可诱发。对于非惯用指令, 由于动词与名词缺乏既有联结, 个体需要临时整合动词与名词以构建动作表征, 可能需要借助视觉通道来产生 SPT 效应。不过实验 1 以物体图片作为视觉线索的载体, 而研究表明物体图片对个体注意力的捕获以及对物体名称的唤醒程度, 均弱于真实物体(Gomez et al., 2018; Snow et al., 2023)。据此, 物体图片可能因视觉线索强度不足, 无法有效促进动词与名词的临时整合, 从而难以诱发 SPT 效应; 个体可能需要与物体发生实际交互, 才能有效构建动作−物体联结, 从而产生 SPT 效应。为检验后一假设, 实验 2 聚焦非惯用指令, 改用真实物体(被试要实际操作物体), 考察物体交互性是否是非惯用指令下 SPT 效应产生的必要条件。如果实验 1 和实验 2 的结果显示“惯用指令的 SPT 效应不依赖物体的呈现, 而非惯用指令的 SPT 效应仅在实物条件下出现”, 则可初步推断: 惯用指令下的 SPT 效应不依赖视觉通道, 非惯用指令下的 SPT 效应则依赖于与物体交互所提供的(动觉以外的)感觉输入。与物体交互涉及视觉与触觉, 实验 2 无法明确哪个感觉通道起作用。实验 3 将分离视觉与触觉的贡献, 在实物呈现前提下, 正交操纵视觉(戴眼罩 vs. 不戴眼罩)与精细触觉(戴手套 vs. 不戴手套), 并设置纯言语编码基线, 检验视觉与精细触觉对非惯用指令下 SPT 效应的独立贡献。通过这三个实验, 本研究将揭示 SPT 效应的感觉通道依赖性如何随指令惯用性而变化。

2　实验 1: 指令惯用性是否调节物体图片呈现对 SPT 效应的影响

实验 1 检验对于由熟悉物体名词构成的言语指令, 指令惯用性是否调节物体视觉线索对 SPT 效应

的影响。如同于战宇(2014)的研究，本实验通过物体图片的呈现与否操纵视觉线索的有无。此外，采用熟悉物体名词，通过动词操控形成惯用指令(如“削铅笔”)与非惯用指令(如“戳铅笔”)。

2.1 方法

2.1.1 被试

基于 Xie 等(2024)报告的交互作用效应量(η_p^2 = 0.082)，采用 G*Power 3.1 进行先验功效分析(α = 0.05, power = 0.95，重复测量方差分析，组内−组间交互)，估算所需总样本量为 56(每组 14 人)。为获得更稳定的效应估计并预留数据剔除空间，我们在数据收集前决定将样本量扩大至每组 50 人，共招募 200 名 18~22 岁的本科生(M_{age} = 20.14 ± 0.93 岁；男性 81 名)。所有被试自述无精神疾病史、视力正常或矫正至正常、右利手且智力正常。按事先制定的标准剔除数据偏离样本均值 3 个标准差以上、或在编码阶段报告走神且回忆得分为 0 的被试。所有被试均未满足剔除标准，全部纳入最终分析。实验结束后，被试收到小额现金报酬。

2.1.2 材料

基于 Wang 等(2022)与 Xie 等(2024)的指令记忆材料库，从中选取全部 16 个物体名词，并补充 13 个自编的常见物品名词，形成 29 个候选名词。为每个候选名词匹配一个符合日常习惯的动词(如为“开关”匹配“按”)，构成惯用指令(如“按开关”)。由于部分动词在匹配中重复出现，为避免动词重复干扰实验结果，剔除重复项。最终保留 20 个名词，并为每个名词分配唯一且不重复的惯用动词，形成 20 条惯用指令。另为这 20 个名词匹配 20 个通常不与之搭配、但动作可执行的动词(如“按镜子”虽语义非常规，但个体能完成“按”这一动作)，构成非惯用指令，组内 20 个动词互不重复。

为检验材料的适用性，邀请 10 名未参与正式实验的本科生(M_{age} = 19.90 ± 0.74 岁；男性 5 名)完成两项评定任务：(1)对 20 个物体名词、20 个惯用动词和 16 个非惯用动词(4 个非惯用动词与惯用动词重复，未纳入评分)分别进行 1~7 级熟悉度评分(1 = 极不熟悉，7 = 极熟悉)；(2)对 40 个动名词短语进行 1~7 级搭配合理性评分(1 = 完全不合理，7 = 完全合理)。结果显示，名词、惯用动词和非惯用动词的熟悉度均在 4~7 分。以 4 分为“较熟悉”标准；所有词汇的评分均不低于 4 分，表明均为被试较熟悉的词汇。在搭配合理性方面，惯用搭配得分(M = 6.67, SD = 0.69，范围 4~7)显著高于非惯用搭配得分(M = 1.38, SD = 0.70，范围 1~3)，$t(199) = 79.78$, $p < 0.001$。

最终采用的 20 条惯用指令为：削铅笔、搬椅子、按开关、端盘子、系鞋带、撑雨伞、挂衣架、刷牙刷、梳梳子、拧毛巾、挤牙膏、搓肥皂、提水桶、照镜子、夹筷子、滑手机、翻书本、擦桌子、戴眼镜、盖被子。

对应的 20 条非惯用指令为：戳铅笔、撕椅子、推开关、掰盘子、敲鞋带、翻雨伞、抠衣架、切牙刷、捏梳子、摸毛巾、梳牙膏、拧肥皂、拍水桶、按镜子、拉筷子、弹手机、刮书本、甩桌子、摇眼镜、转被子。

2.1.3 设计

采用 2 (指令惯用性：惯用 vs. 非惯用) × 2 (物体图片呈现：呈现 vs.不呈现) × 2 (编码方式：言语 vs. 动作)混合设计。编码方式为被试内变量，其余为被试间变量。因变量指标为每种实验条件正确回忆的指令数量。

2.1.4 程序

实验流程图见图 1。将被试随机分配至四组之一：(1)惯用−图片呈现组；(2)惯用−图片不呈现组；(3)非惯用−图片呈现组；(4)非惯用−图片不呈现组。“惯用”指语义合理的指令(如“削铅笔”)；“非惯用”指语义非常规但动作可执行的指令(如“戳铅笔”)。所有被试需完成动作编码与言语编码条件任务，每种条件包含 10 条指令的学习与回忆任务。两种条件呈现顺序在被试间平衡。具体任务流程如下：

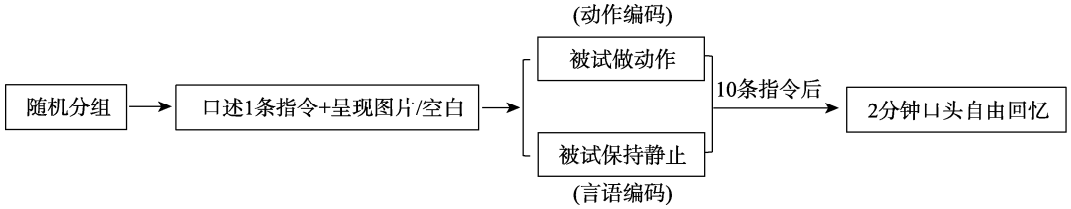


图 1 实验 1 流程图。随机分组指将被试随机分配到惯用−图片呈现组、惯用−图片不呈现组、非惯用−图片呈现组、非惯用−图片不呈现组。惯用−图片呈现组呈现惯用指令和物体图片，惯用−图片不呈现组呈现惯用指令和空白纸，非惯用−图片呈现组呈现非惯用指令和物体图片，非惯用−图片不呈现组呈现非惯用指令和空白纸。

(1)动作编码条件: 主试口述一条指令(持续约 3 秒)的同时, 图片呈现组的被试会看到指令中名词所对应物体的彩色实物图片(由主试手持展示 7 秒), 图片不呈现组则看到一张空白 A4 纸(由主试手持展示 7 秒)。在此期间, 被试需根据听到的指令做出相应动作(例如, 听到“削铅笔”时, 用手做出削铅笔的动作)。10 条指令呈现完毕后, 被试需在 2 分钟内口述所听到的 10 条指令, 不限顺序。

(2)言语编码条件: 同样呈现彩色实物图片或白纸(各 7 秒), 但被试保持手部静止, 仅听指令。10 条指令呈现完毕后, 被试进行 2 分钟口头自由回忆。

对被试的口头回忆进行录音, 由一名不了解实验目的的非心理学专业学生独立评分。评分标准为: 被试口述的指令须与学习阶段呈现的指令一致, 动词和名词均正确即计为正确(如学习阶段为“削铅笔”, 被试须说出“削铅笔”), 对顺序不做要求; 仅说出名词、仅说出动词, 或动词与名词不匹配, 均计为错误。

2.2　结果

四组被试在年龄($F(3, 196) = 0.78$, $p = 0.504$)和性别比例($\chi^2(3) = 3.88$, $p = 0.275$)上无显著差异。对所有被试的正确回忆数进行三因素重复测量方差分析, 描述性统计结果见图 2。指令惯用性主效应显著, $F(1, 196) = 185.42$, $p < 0.001$, $\eta_p^2 = 0.49$。物体图片呈现主效应显著, $F(1, 196) = 144.96$, $p < 0.001$, $\eta_p^2 = 0.43$。编码方式主效应显著, $F(1, 196) = 42.47$, $p < 0.001$, $\eta_p^2 = 0.18$。指令惯用性与物体图片呈现的交互作用显著, $F(1, 196) = 22.67$, $p < 0.001$, $\eta_p^2 = 0.10$。指令惯用性与编码方式的交互作用显著, $F(1, 196) = 21.93$, $p < 0.001$, $\eta_p^2 = 0.10$。物体图片呈现与编码方式的交互作用不显著, $F(1, 196) = 0.39$, $p = 0.531$, $\eta_p^2 = 0.002$。指令惯用性、物体图片呈现与编码方式的三阶交互作用不显著, $F(1, 196) = 0.67$, $p = 0.413$。

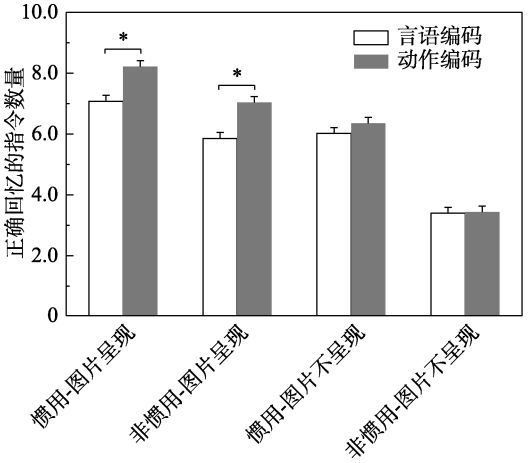


图 2　实验 1 中各条件正确回忆指令数量的均值与标准误

注: * $p < 0.05$, 下同

根据统计学规范, 基于先验理论假设的简单效应分析不依赖于整体交互效应的显著性(Clark-Carter, 2018; Howell, 2013)。鉴于本研究对惯用指令与非惯用指令下的 SPT 效应模式有明确的理论预期, 即预期惯用指令在有无图片条件下均出现 SPT 效应, 而非惯用指令则均不出现, 我们对四种条件(指令惯用性×物体图片呈现)下动作编码与言语编码的差异进行了 Bonferroni 校正的简单效应分析。结果显示: 对于惯用指令, 无论图片是否呈现, 动作编码回忆数均多于言语编码回忆数(图片呈现时: $M_{动作} = 8.22 \pm 0.98$ vs. $M_{言语} = 7.08 \pm 1.37$, $p < 0.001$, 95% CI = [0.73, 1.55]; 无图片时: $M_{动作} = 7.04 \pm 1.18$ vs. $M_{言语} = 5.87 \pm 1.20$, $p < 0.001$, 95% CI = [0.77, 1.59]), 即 SPT 效应都存在。对于非惯用指令, 无论图片是否呈现, 动作编码与言语编码回忆数均无差异(图片呈现时: $M_{动作} = 6.36 \pm 1.63$ vs. $M_{言语} = 6.02 \pm 1.41$, $p = 0.102$, 95% CI = [−0.07, 0.75]; 无图片时: $M_{动作} = 3.44 \pm 1.64$ vs. $M_{言语} = 3.40 \pm 1.49$, $p = 0.847$, 95% CI = [−0.37, 0.45]), 即 SPT 效应都不存在。进一步比较惯用指令在有、无图片条件下的 SPT 效应量(即动作编码条件成绩−言语编码条件成绩; 也见 Allen et al., 2019; Cohen, 1981; Xie et al., 2024), 两者无显著差异, $M_{有图片} = 1.14 \pm 1.37$ vs. $M_{无图片} = 1.18 \pm 1.47$, $t(98) = -0.14$, $p = 0.888$, Cohen's $d = 0.03$, 95% CI = [−0.60, 0.52]。

2.3　讨论

实验 1 发现物体图片的呈现与否对 SPT 效应的影响不随指令惯用性而异。具体而言, 对于惯用指令, SPT 效应在有/无物体图片条件下均存在。这表明, 当指令为日常生活中常规的动宾搭配(如“削铅笔”)时, 外部视觉线索并非必要, 动觉输入足以支持有效的记忆加工。这一结果与于战宇(2014)的发现一致。相反, 对于非惯用指令(如“戳铅笔”), 无论物体图片是否呈现, SPT 效应均未出现。这一结果提示, 当指令的动词与名词缺乏既有联结时, 仅依靠无实物的模拟动作可能难以建立稳定的动作−言语整合表征, 从而无法促进记忆。

需要注意的是, “非惯用指令在有无物体图片呈现时均无 SPT 效应”并不等同于“视觉线索无效”。

若动作编码本身无法启动有效的记忆加工(如因缺乏对物体功能或操作方式的感知), 视觉输入的潜在增益将无法显现。换言之, 效应缺失可能源于基线加工失败, 而非视觉线索无效。实验 1 以图片为视觉线索, 其信息强度(无三维结构、无动作−物体交互)可能不足以支持非惯用指令的在线动作规划。此外, 实验 1 采用被试间设计检验物体图片的作用, 个体差异可能增大误差方差, 相对降低统计效力, 限制对微弱效应的检测。实验 2 将物体图片替换为真实物体, 允许被试进行真实的动作−物体交互, 并在被试内设计下检验非惯用指令能否通过动作编码诱发 SPT 效应。

3 实验2: 实物呈现对非惯用指令下SPT效应的影响

实验 1 发现, 非惯用指令在有无物体图片条件下均未出现 SPT 效应。为检验该效应缺失是否源于视觉线索不充足, 实验 2 将图片替换为可交互的真实物体, 探索当提供更丰富的感觉输入(含视觉、触觉及与动作相关的物理信息), 非惯用指令下动作编码能否诱发 SPT 效应。

3.1 方法

3.1.1 被试

基于 Xie 等(2024)报告的交互作用效应量(η_p^2 = 0.082), 采用 G*Power 3.1 进行先验功效分析(α = 0.05, power = 0.95, 重复测量方差分析, 组内因子), 所需样本量为 26。为获得更稳定的效应估计并预留数据剔除空间, 我们在数据收集前决定将样本量扩大至 64 人, 共招募 64 名本科生(M_{age} = 19.59 ± 1.07 岁; 男性 23 名)。按照实验 1 的剔除标准,实验 2 剔除了 2 名在编码阶段报告走神且回忆得分为 0 的被试, 以及 4 名存在极端数据的被试(在 4 个条件共 256 个观测值中, 任一观测值偏离该条件均值 3 个标准差, 即剔除该被试全部数据),最终纳入 58 名被试的数据(M_{age} = 19.55 ± 1.06 岁; 男性 20 名)。所有被试自述无精神疾病史、视力正常或矫正至正常、右利手且智力正常。

3.1.2 材料

鉴于实验 1 所用部分物体(如被子)难以在桌面实验环境实际操作, 本实验未沿用其实物材料。参照实验 1 的搭配逻辑, 从 Wang 等(2022)与 Xie 等(2024)的材料库中选取 13 个适合桌面操作的高熟悉度物品名词(剪刀、钥匙、铅笔、梳子、毛巾、杯子、纸巾、雨伞、勺子、筷子、镜子、口罩、牙刷), 结合 10 个高频动词(转、摸、推、摇、敲、翻、拉、按、切、捏)生成 130 条非惯用指令(如“切毛巾”)。

邀请参与实验 1 评定的 10 名本科生(5 男 5 女)对材料进行两项评定: (1)对 13 个名词与 10 个动词进行 1~7 级熟悉度评分(1 = 极不熟悉, 7 = 极熟悉); (2)对 130 条指令进行 1~7 级搭配合理性评分(1 = 完全不合理, 7 = 完全合理)。结果显示, 所有名词和动词的熟悉度均较高(范围 5~7); 指令的名词与动词搭配合理性均较低(范围 1~3, 低于 7 点量表的中点 4 分), 具有明显语义非常规性。从中选出 40 条非惯用指令。筛选标准为: 动词与名词在各序列内不重复, 以控制材料间的相互干扰。入选的 40 条指令组成 4 个序列, 每序列 10 条指令。所有指令均可由被试在桌面环境中单手完成。

3.1.3 设计

采用 2 (物体呈现: 有 vs.无) × 2 (编码方式: 言语 vs.动作)被试内设计。因变量指标为每个条件下正确回忆的指令数量。

3.1.4 程序

实验流程图见图 3。每位被试需完成 4 种实验条件, 条件顺序在被试间平衡。每种条件包含 10 条非惯用指令的编码与回忆阶段。编码阶段各条件任务如下:

(1)有物体动作编码条件: 主试在口述一条指令(约 3 秒)的同时, 将对应物体置于桌面中央(呈现 7 秒); 被试需在此期间对物体执行与动词相符的动

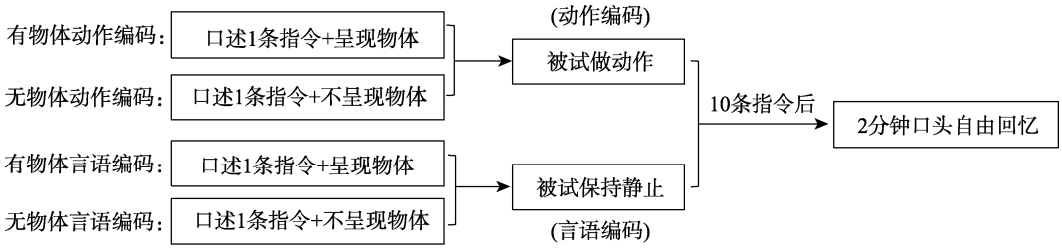


图 3 实验 2 流程图。每个被试要完成 4 个条件任务: 有物体动作编码、无物体动作编码、有物体言语编码、无物体言语编码。

作(如听到“切毛巾”, 用手对毛巾做出“切”的动作)。(2)无物体动作编码条件: 桌面无物体; 被试听到指令时, 需在此期间无实物做出相应动作。(3)有物体言语编码条件: 主试在口述指令(约 3 秒)的同时, 将对应物体置于桌面中央(7 秒); 被试需在此期间保持手部静止, 仅听指令, 不得做出任何动作。(4)无物体言语编码条件: 桌面无物体; 被试听到指令(3 秒)时, 要保持手部静止, 仅听指令。

所有条件下, 10 条指令编码完成后, 被试在 2 分钟内尽可能口头回忆这 10 条指令, 不限顺序。对被试的口头回忆进行录音, 评分方式同实验 1。

3.2 结果

对每个条件的正确回忆数进行两因素重复测量方差分析, 描述性统计结果见图 4。编码方式主效应显著, $F(1, 57) = 28.29$, $p < 0.001$, $\eta_p^2 = 0.33$。物体呈现主效应显著, $F(1, 57) = 247.15$, $p < 0.001$, $\eta_p^2 = 0.81$。物体呈现与编码方式的交互作用显著, $F(1, 57) = 16.65$, $p < 0.001$, $\eta_p^2 = 0.23$。Bonferroni 校正的简单效应分析显示, 有物体呈现时, 动作编码条件的回忆数($M = 6.69$, $SD = 1.52$)多于言语编码条件的回忆数($M = 5.03$, $SD = 1.77$), $p < 0.001$, 95% CI = [1.09, 2.22]。无物体呈现时, 动作编码条件的回忆数($M = 3.09$, $SD = 1.48$)与言语条件的回忆数($M = 2.81$, $SD = 1.28$)无差异, $p = 0.188$, 95% CI = [−0.14, 0.69]。动作编码和言语编码时, 有物体条件的回忆数均多于无物体条件的回忆数, ps < 0.001, 95% CI 均不包含 0。

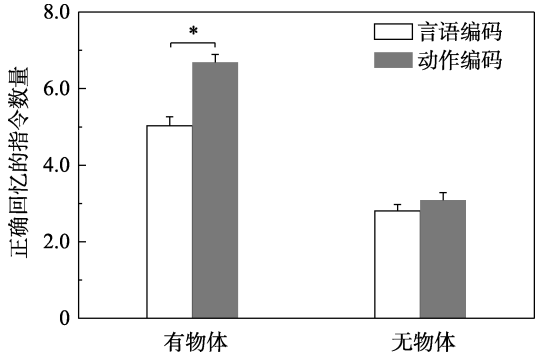


图 4　实验 2 中各条件正确回忆指令数量的均值与标准误

3.3 讨论

实验 2 发现, 非惯用指令仅在有物体呈现时出现 SPT 效应。这与实验 1 中图片条件下无 SPT 效应的结果形成对比, 提示与真实物体交互可能是非惯用指令产生 SPT 效应的关键。不过, 实验 1 与实验 2 在实验设计、材料上存在差异。为排除这些差异对非惯用指令下视觉线索效应的潜在混淆, 我们补充了一个与实验 2 同结构的图片实验(被试为 26 名与实验 2 被试来自同一学校的本科生, 10 名男生, 采用被试内设计, 材料同实验 2, 以物体图片替代实物, 被试对图片执行模拟动作)。结果显示, 物体图片呈现与编码方式的交互作用不显著($F(1, 25) = 0.84$, $p = 0.368$), 编码方式主效应($F(1, 25) = 1.60$, $p = 0.218$)、物体图片呈现主效应($F(1, 25) = 1.74$, $p = 0.199$)亦不显著(有图片动作编码的回忆数: $M = 3.92$, $SD = 1.32$; 有图片言语编码的回忆数: $M = 4.08$, $SD = 1.32$; 无图片动作编码的回忆数: $M = 3.38$, $SD = 1.53$; 无图片言语编码的回忆数: $M = 3.96$, $SD = 1.69$)。这表明, 即使在统计效力更高的被试内设计中, 图片条件依然无法使非惯用指令产生 SPT 效应。

综合实验 1、实验 2 及补充实验的结果, 对于非惯用指令, SPT 效应的出现依赖可交互的真实物体在场。当仅有物体图片或无物体时, 单纯的动作模拟可能因缺乏对物体物理属性(如形状、质地和可操作性)的直接感知, 难以有效整合动作与言语信息; 而真实物体通过同步提供视觉、触觉与动作反馈, 使动作编码得以成功启动。

从多感觉通道的角度看, 实验 1 的结果表明视觉通道不是惯用指令下 SPT 效应产生所必需的。而实验 1、实验 2 及补充实验的结果共同表明, 对于非惯用指令, 仅凭动觉不足以产生 SPT 效应, 尚需动觉以外的其他感觉通道参与。然而, 真实物体同时激活视觉与触觉, 当前结果无法确定起关键作用的是视觉、触觉还是二者兼有。为此, 实验 3 将采用正交设计, 操纵视觉与精细触觉输入的有无, 检验非惯用指令下 SPT 效应的产生依赖于哪些感觉通道。

4 实验 3: 视觉与触觉通道在 SPT 效应中的独立贡献

实验 3 通过操纵视觉与精细触觉的可用性, 分离二者对非惯用指令下 SPT 效应的贡献。若某一感觉输入具有关键作用, 则其剥夺将削弱或消除 SPT 效应; 若二者协同作用, 则仅在视觉与精细触觉均可用时效应最大。

4.1 方法

4.1.1 被试

基于 Xie 等(2024)报告的交互作用效应量($\eta_p^2 =$

0.082), 采用 G*Power 3.1 进行先验功效分析(α = 0.05, power = 0.95, 重复测量方差分析, 组内因子), 所需样本量为 23。为获得更稳定的效应估计并预留数据剔除空间, 我们在数据收集前决定将样本量扩大至 54 人, 招募 54 名本科生(M_{age} = 19.80 ± 1.17 岁; 男性 25 名)。剔除标准同实验 1。所有被试均未被剔除, 全部纳入最终分析。所有被试自述无神经或精神系统疾病史、视力正常或矫正至正常, 且为右利手。

4.1.2 材料

实验材料同实验 2。10 条非惯用指令构成一个序列, 每个序列内动词与名词不重复。共编制 5 个序列, 分别用于 5 种实验条件。

4.1.3 设计

采用单因素被试内设计, 自变量为编码方式, 包含 5 个水平: 有视觉有触觉动作编码、有视觉无精细触觉动作编码、无视觉有触觉动作编码、无视觉无精细触觉动作编码、言语编码。因变量指标为每个条件下正确回忆的指令数量。

4.1.4 程序

实验流程图见图 5。每位被试需完成 5 种实验条件, 条件顺序在被试间平衡。每种条件包含 10 个非惯用指令的编码与回忆阶段。编码阶段各条件任务如下:

(1)有视觉有触觉动作编码条件: 主试口述一条指令(约 3 秒)的同时, 将对应物体置于桌面中央(7 秒); 被试需在此期间对物体执行与动词相符的动作(如听到“切毛巾”时, 用手对毛巾做出“切”的动作)。(2)有视觉无精细触觉动作编码条件: 被试事先戴好厚棉手套(阻断纹理与温度感知, 但保留粗略形状与阻力); 主试口述指令(约 3 秒)的同时, 将对应物体置于桌面中央(7 秒); 被试需在此期间对物体执行与动词相符的动作。(3)无视觉有触觉动作编码条件: 被试事先戴好全遮光硅胶眼罩(确保无光泄漏); 主试口述指令(约 3 秒)的同时, 将对应物体置于桌面中央(7 秒); 被试需在此期间对该物体执行与动词相符的动作。(4)无视觉无精细触觉动作编码条件: 被试事先戴好厚棉手套(阻断纹理与温度感知, 但保留粗略形状与阻力)和全遮光硅胶眼罩(确保无光泄漏); 主试口述指令(约 3 秒)的同时, 将对应物体置于桌面中央(7 秒); 被试需在此期间对该物体执行与动词相符的动作。(5)言语编码条件: 桌面无物体; 被试听到指令(3 秒)时, 只能听指令, 不能做出任何动作。

所有条件下, 10 条指令编码完成后, 被试在 2 分钟内尽可能口述这 10 条指令, 不限顺序。对被试的口头回忆进行录音, 评分方式同实验 1。

4.2 结果

对每种条件的正确回忆数进行单因素重复测量方差分析, 描述性统计结果见图 6。编码方式主效应显著, $F(4, 212) = 19.26$, $p < 0.001$, $\eta_p^2 = 0.27$。Bonferroni 校正的事后多重比较显示, 有视觉有触觉动作编码条件的回忆数(M = 6.89, SD = 1.69)与有视觉无精细触觉动作编码条件的回忆数(M = 6.83, SD = 1.66)无差异($p > 0.999$), 二者均多于无视觉有触觉动作编码(M = 5.13, SD = 1.94)、无视觉无精细触觉动作编码(M = 5.24, SD = 1.76)和言语编码条件的回忆数(M = 4.78, SD = 1.93) (ps < 0.001), 后三个条件的回忆数无差异(ps > 0.999)。这一结果说明只有两个有视觉条件(即有视觉有触觉动作编码条件与有视觉无精细触觉动作编码条件)出现 SPT 效应。

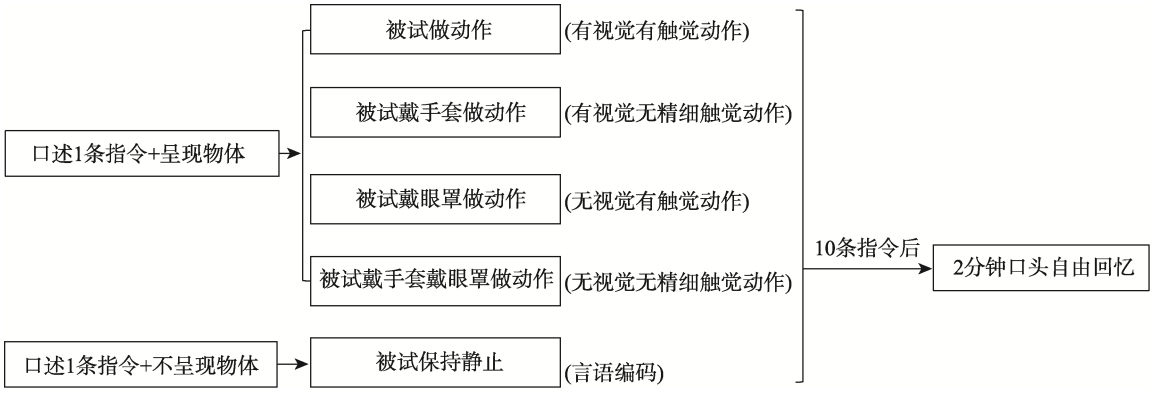


图 5 实验 3 流程图。每个被试要完成 5 个条件任务: 有视觉有触觉动作编码、有视觉无精细触觉动作编码、无视觉有触觉动作编码、无视觉无精细触觉动作编码条件、言语编码。其中, 无视觉时被试需要戴眼罩, 无精细触觉时被试需要戴手套。

为进一步检验SPT效应的存在性与强度, 实验3 计算各动作编码条件相对于言语基线的增益值(即动作编码条件成绩−言语编码条件成绩), 作为SPT效应量指标。对四组 SPT 效应量进行单因素重复测量方差分析, 结果显示主效应显著, $F(3, 159) = 19.25$, $p < 0.001$, $\eta_p^2 = 0.27$。Bonferroni 校正的事后多重比较显示, 有视觉条件下(无论触觉是否可用), SPT 效应量均较大(有视觉有触觉动作编码条件: $M = 2.11$, $SD = 2.29$; 有视觉无精细触觉动作编码条件: $M = 2.06$, $SD = 2.39$), 且二者无差异($p > 0.999$)。无视觉条件(无论触觉是否可用)的 SPT 效应量较小(无视觉有触觉动作编码条件: $M = 0.35$, $SD = 2.75$; 无视觉无精细触觉动作编码条件: $M = 0.46$, $SD = 2.60$), 且二者无差异($p > 0.999$)。有视觉有触觉动作编码和有视觉无精细触觉动作编码条件的 SPT 效应量均大于无视觉有触觉动作编码和无视觉无精细触觉动作编码条件的 SPT 效应量(ps < 0.001)。

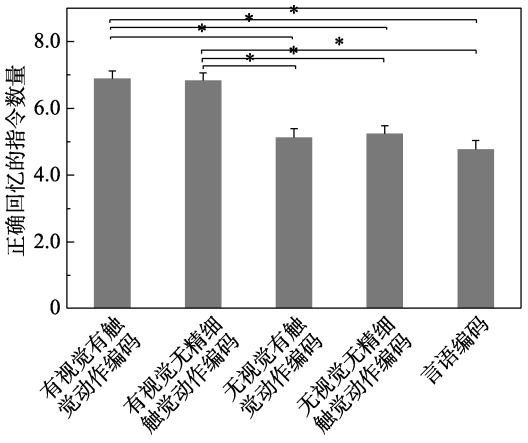


图 6　实验 3 中各条件正确回忆指令数量的均值与标准误

4.3　讨论

实验 3 通过操纵视觉与精细触觉的可用性, 分离二者在非惯用指令 SPT 效应中的贡献。结果发现, 无论视觉是否可用, 触觉精细信息(如纹理与温度)的剥夺均不影响 SPT 效应的出现, 表明触觉精细信息并非 SPT 效应所必需。相比之下, 视觉对 SPT 效应具有关键作用: 仅当视觉通道开放时, SPT 效应量达到较高水平(均值≈2.0); 而视觉剥夺使效应量显著降低(均值< 0.5)。综合三项实验可知, 对于由熟悉物体名词构成的非惯用指令, SPT 效应的产生需要被试与真实物体进行交互(实验 2 vs.实验 1、补充实验), 且该交互中起关键作用的是视觉通道, 而非精细触觉输入(实验 3)。

除了分离视觉与精细触觉的贡献, 实验 3 的结果也有助于回应实验 2 的一个潜在替代性解释: 真实物体相较于图片有更强的可供性与多模态感知(Gomez et al., 2018; Snow et al., 2023), 因此实验 2 中非惯用指令在真实物体条件下的 SPT 效应可能并非源于动作执行, 而是源于真实物体的这些属性。若 SPT 效应完全由真实物体的可供性驱动, 在有实物条件下, 动作编码与言语编码的成绩应当无差异, 因为两种条件下被试都看到了实物, 可供性的激活程度应当相同。但实验 2 的数据显示, 有物体动作编码条件的回忆数($M = 6.69$)显著高于有物体言语编码条件($M = 5.03$)。这说明, 动作执行在可供性基础上额外促进了记忆。关于多模态感知, 实验 3 中有视觉有触觉与有视觉无精细触觉两种条件均使用真实物体, 被试均执行动作, 仅触觉精细信息的可用性不同。两种条件下的 SPT 效应无显著差异。若真实物体的优势完全来自其多模态感知(含触觉成分), 减弱触觉信息应当削弱效应; 但结果并非如此。综合上述两点分析, 实验数据不支持将可供性或多模态感知单独作为实验 2 中真实物体下 SPT 效应的解释, 动作执行本身在该效应中不可或缺。关于动作执行中具体是哪些成分在发挥作用, 实验 2 无法直接回答。实验 3 虽然对动作执行中的感觉成分做了分离, 但鉴于实验 2 并未操纵视觉通道, 将实验 3 的发现回溯应用于实验 2 属于逻辑推演, 其解释力存在局限。未来研究可直接分离视觉与动觉, 如增设“呈现实物但仅进行模拟动作”的条件, 检验是视觉−动觉协同还是其他与真实物体交互相关的加工导致了实验 2 的结果。

5　总讨论

本研究通过三个实验考察 SPT 效应所依赖的感觉通道如何随指令惯用性而变化, 聚焦两个问题: (1)对于由熟悉物体名词(如“铅笔”)构成的指令, 惯用搭配(如“削铅笔”)与非惯用搭配(如“戳铅笔”)在动作编码下诱发的 SPT 效应是否表现出不同的感觉通道依赖性? (2)精细触觉输入是否对 SPT 效应具有独立贡献? 实验 1 发现, 对于惯用指令, 无论是否呈现物体图片, SPT 效应均存在, 表明当动词与名词有既有联结时, 动觉输入即可支持SPT效应的产生, 无需外部视觉线索; 而对非惯用指令, 无论有无物体图片, SPT 效应均未出现, 说明动词与名词缺乏既有联结时, 仅动作执行难以产生记忆优

势。实验 2 聚焦非惯用指令, 发现在真实、可交互物体在场时, SPT 效应出现; 其补充的图片条件(被试内设计, 材料同实验 2)则未观察到 SPT 效应, 排除了实验设计差异的混淆, 进一步确认了物体交互性对非惯用指令编码的关键作用。由实验 1 和 2 可知, 惯用指令的 SPT 效应不依赖物体视觉线索, 而非惯用指令则依赖。实验 3 通过正交操纵视觉(眼罩)与精细触觉(手套)输入, 发现对于非惯用指令, SPT 效应的存在依赖视觉通道(有视觉时 SPT 效应量较大, 无视觉时 SPT 效应量明显减小), 而不依赖于触觉的精细信息(如纹理、温度)。综合所有实验, SPT 效应的感觉通道依赖模式随指令惯用性而异: 惯用指令的 SPT 效应可由动觉独立驱动; 非惯用指令的 SPT 效应则需动觉与视觉协同作用, 触觉的精细信息在此过程中并非必需。

为何非惯用指令的 SPT 效应依赖视觉与动觉的协同, 而精细触觉几乎不起作用? 根据多通道加工理论, 执行动作在言语编码的基础上增加了额外的感觉通道, 从而加深了指令的记忆痕迹, 使其更易被提取(Bäckman & Nilsson, 1984)。对于惯用指令(如“削铅笔”), 动词与名词的联结事先存在, 执行动作可直接加深整个短语的记忆痕迹。然而, 对于非惯用指令(如“拉筷子”), 动词与名词缺乏既有联结。执行动作虽然可以加深动词(如“拉”)的痕迹, 但名词(如“筷子”)的痕迹难以通过动作本身得到增强, 可能需要借助视觉信息。有研究发现, 视觉系统可以快速获取物体的空间特征和功能信息(Contier et al., 2024; Ruotolo et al., 2020)。由此推测, 在编码阶段, 视觉可以帮助个体识别筷子的长轴方向、操作端点及功能表面等, 形成“如何操作该物体”的指引(如识别“拉”的方向和方式)。当个体依据这一指引实际执行动作时, 视觉信息与动觉信息会相互整合, 名词的痕迹可能因视觉提供的物体信息而增强, 动词的痕迹可能因动觉执行而加深, 整个动名词短语的记忆痕迹可能因整合而得到加深。若视觉被阻断, 则可能无法形成明确的操作指引(如不知从哪拉、朝哪个方向拉), 名词的痕迹因缺乏物体信息的支撑而难以增强, 记忆效果随之削弱。不同于视觉, 精细触觉在痕迹加深过程中可能既非关键也非充分。一方面, 研究表明在视觉可用时, 大脑基于长期经验已建立高效的“视觉→触觉”预测(Li et al., 2019)。当实际触觉输入(如物体的硬度、纹理)与视觉预测一致时, 该信号可能被视为冗余而被抑制, 以提升加工效率。另一方面, 即便在视觉缺失情境下, 触觉虽能传递局部物理属性, 却无法提供物体整体的空间朝向或操作轴线, 因而可能难以指导动作参数的设定(如无法判断“拉”的起始点或方向)。针对盲人的研究间接支持了这一推测: 即使是触觉功能因长期代偿而远超常人的盲人, 其 SPT 效应仍弱于视力正常者(Kormi-Nouri, 1995), 反向表明, 触觉即使在最优条件下也难以替代视觉在加深记忆痕迹中的核心作用。考虑到上述解释只是基于多通道加工理论对实验 3 结果的推测, 未来研究可通过检验“名词痕迹的增强是否必需视觉系统的参与”及“视觉系统是否足以独立指导动作参数的设定”来提供更确凿的证据。

总之, 本研究揭示, 感觉通道在 SPT 效应中的贡献并非恒定, 而是随指令惯用性而不同。这些发现有助于完善经典的多通道加工理论(Bäckman & Nilsson, 1984)。该理论强调多通道信息可带来更丰富的编码(rich encoding)以提升记忆, 但未明确哪些通道对 SPT 效应起关键作用、哪些作用有限。在本研究实验 1 中, 惯用指令在无图片条件(仅动觉通道)与有图片条件(动觉与视觉双通道)产生了无差异的 SPT 效应, 即双通道条件下并未观察到更强的效应。这一结果与多通道理论“通道越丰富编码越强”的预期不同。这意味着, 对于具有既有联结的惯用指令(如“削铅笔”), 其 SPT 效应并不依赖通道数量的增加。而对于缺乏既有联结的非惯用指令(如“拉筷子”), 实验 2 和实验 3 的结果呈现出另一种模式: 单凭动觉无法产生 SPT 效应, 需视觉与动觉协同, 但并非所有通道的加入都有效, 触觉精细信息的加入并未进一步增强效应。由此可知, 对于非惯用指令, 多通道理论能够解释为何需要多通道协同, 但无法说明哪些通道的加入才是关键的。综合三项实验结果, 通道的贡献并非简单地取决于通道数量, 而是与指令是否具有既有联结有关: 指令具有既有联结时, 动觉通道即可; 缺乏既有联结时, 则需要额外通道辅助。因此, 多通道理论需要从“多通道带来更丰富编码”的认识, 走向“通道贡献因经验而异”的更为分化的认识。

在实践层面, 本研究为“动作编码促进记忆”策略提供了精细化应用指南。当个体需处理动词与名词之间缺乏既有联结的非惯用指令(如操作新设备、理解陌生流程)时, 仅靠肢体动作难以有效提升记忆; 必须同步提供可见物体(实物或高保真视觉呈现), 以建立视觉–空间锚点, 支撑临时联结的在线构建。值得注意的是, 精细触觉并非必要条件,

在佩戴手套等触觉受限情境下, 只要视觉通道开放, 自我执行动作仍能促进记忆。这一发现具有重要应用价值: 在卫生敏感场景(如实验室、厨房)或存在触觉敏感/回避特征的群体中, 只需确保视觉信息可见且物体可交互, 无需依赖精细触觉反馈, 即可有效实施动作编码干预, 提高指令记忆。

本研究亦存在局限。首先, 本研究通过动词操控形成惯用指令与非惯用指令, 但这一操纵同时伴随语义合理性、动作自然性等多个维度的差异。虽然这些维度在实验 1 中惯用指令内部和非惯用指令内部各自是恒定的, 因而不影响两类指令内部有无图片条件对比的结论, 但可能影响实验 1 结果的推广性。此外, 非惯用指令内部也存在一定异质性, 部分指令如“摸毛巾”在特定情境下可发生, 而“切牙刷”更接近语义异常。这一内部异质性在有无图片条件之间通过随机分配而平衡, 因此不影响该对比的统计结论, 但可能限制了“非惯用”这一概念的统一性。未来研究可进一步区分非惯用指令的子类型, 系统检验这些维度对 SPT 效应的影响。第二, 本研究将物体熟悉度作为恒定变量(三项实验均使用高熟悉度日常物品), 考察指令惯用性的独立作用, 因此核心结论限定于熟悉物体。对于不熟悉物体, 惯用性是否发挥相同的作用尚待检验, 未来研究可在本研究基础上操纵物体熟悉度, 系统检验熟悉度是否调节惯用性在多通道编码中的作用。第三, 实验 1 与实验 2 在物体呈现方式(图片 vs. 实物)与操作方式(模拟动作 vs. 真实交互)上同步变化, 无法完全分解实物条件下视觉强度与物体−动作真实交互各自的贡献。虽然不影响本研究关于“动觉与视觉交互是非惯用指令下产生 SPT 效应的关键”及“精细触觉并非必要”的核心结论, 但限制了在此基础上对视觉与操作各自作用的更精细分离。未来研究可增设“呈现实物但仅做模拟动作”的条件以进一步澄清。第四, 本研究聚焦视觉与触觉, 物体呈现可能同时诱发嗅觉、听觉(如物体碰撞声)等其他感觉输入, 其潜在贡献尚未被检验。未来研究可借助气味控制室、消音环境等范式, 拓展多通道交互的考察维度。最后, 本研究揭示了视觉与动觉对于 SPT 效应的必要性, 但受限于行为实验的设计, 未能刻画二者在实时加工中的协同过程。未来工作可结合眼动追踪、动作运动学分析或计算建模, 精细刻画视觉注意分配与动作执行在时间上的动态耦合, 进一步揭示视觉与动觉协同促进记忆的内在过程。

6　结论

(1)惯用指令与非惯用指令的 SPT 效应具有不同的感觉通道依赖性: 惯用指令可由动觉独立驱动, 非惯用指令则需要视觉与动觉协同。

(2)触觉精细信息对 SPT 效应无独立贡献。

参 考 文 献

Allen, R. J., Hill, L. J. B., Eddy, L. H., & Waterman, A. H. (2019). Exploring the effects of demonstration and enactment in facilitating recall of instructions in working memory. *Memory & Cognition, 48*(3), 400−410. https://doi.org/10.3758/s13421-019-00978-6

Bäckman, L., & Nilsson, L. G. (1984). Aging effects in free recall: An exception to the rule. *Human Learning: Journal of Practical Research & Applications, 3*(1), 53−69.

Clark-Carter, D. (2018). *Quantitative psychological research: The complete student's companion* (4th ed., pp. 231−258). Routledge.

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology, 22*(1), 267−281. https://doi.org/10.1111/j.1467-9450.1981.tb00402.x

Contier, O., Baker, C., & Hebart, M. (2024). Distributed representations of behaviour-derived object dimensions in the human visual system. *Nature Human Behaviour, 8*, 2179−2193. https://doi.org/10.1038/s41562-024-01980-y

Engelkamp, J. (1996). Organisation and recall in verbal tasks and in subject-performed tasks. *European Journal of Cognitive Psychology, 8*(3), 257−274.

Engelkamp, J., & Zimmer, H. D. (1984). Motor programme information as a separable memory unit. *Psychological Research, 46*(3), 283−299. https://doi.org/10.1007/BF00308889

Gomez, M. A., Skiba, R. M., & Snow, J. C. (2018). Graspable objects grab attention more than images do. *Psychological Science, 29*(2), 206−218. https://doi.org/10.1177/0956797617730599

Howell, D. C. (2013). *Statistical Methods for Psychology* (8th ed.). Wadsworth Cengage Learning

Kormi-nouri, R. (1995). The nature of memory for action events: An episodic integration view. *European Journal of Cognitive Psychology, 7*(4), 337−363.

Kormi-Nouri, R. (2000). The role of movement and object in action memory: A comparative study between blind, blindfolded and sighted subjects. *Scandinavian Journal of Psychology, 41*(1), 71−75.

Li, Y., Zhu, J., Tedrake, R., & Torralba, A. (2019). *Connecting touch and vision via cross-modal prediction*. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 10601−10610. https://doi.org/10.1109/cvpr.2019.01086

Ma, J. L., Wang, L., & Li, Y. (2025). Imagery processing is necessary for the subject-performed task effect: Evidence from event-related potentials. *Learning and Motivation*, *89*, 102091. https://doi.org/10.1016/j.lmot.2024.102091

Mecklenbräuker, S., Steffens, M. C., Jelenec, P., & Goergens, N. K. (2011). Interactive context integration in children? Evidence from an action memory study. *Journal of Experimental Child Psychology, 108*(4), 747−761. https://doi.org/10.1016/j.jecp.2010.12.002

Nyberg, L., Nilsson, L. G., & Bäckman, L. (1991). A component analysis of action events. *Psychological Research, 53*(3), 219−225.

Ruotolo, F., Kalénine, S., & Bartolo, A. (2020). Activation of manipulation and function knowledge during visual search for objects. *Journal of Experimental Psychology: Human Perception and Performance, 46*(1), 66−90. https://doi.org/10.1037/xhp0000696

Snow, J. C., Gomez, M. A., & Compton, M. T. (2023). Human memory for real-world solid objects is not predicted by responses to image displays. *Journal of Experimental Psychology: General, 152*(10), 2703−2712. https://doi.org/10.1037/xge0001387

Wang, L. J., & Li, G. Z. (2014). A review on the theories of subject performed task effect in action memory. *Journal of Psychological Science, 37*(4), 998−1001.

[王丽娟, 李广政. (2014). 动作记忆 SPT 效应的理论探讨. *心理科学, 37*(4), 998−1001.]

Wang, L. J., & Ma, J. L. (2023). Imagery encoding in action memory of art students and non-art students. *Psychological Development and Education, 39*(4), 457−464.

[王丽娟, 马嘉琳. (2023). 美术生和非美术生的动作记忆表象编码. *心理发展与教育, 39*(4), 457−464.]

Wang, L. J., Xie, T. T., Ma, H., Xu, M., & Xie, X. C. (2022). Subject-performed task effect on working memory performance in children with autism spectrum disorder. *Autism Research, 15*(9), 1698−1709. https://doi.org/10.1002/aur.2710

Xie, T. T., Ma, H., Wang, L. J., & Du, Y. F. (2024). Can enactment and motor imagery improve working memory for instructions in children with autism spectrum disorder and children with intellectual disability? *Journal of Autism and Developmental Disorders, 54*(1), 131−142. https://doi.org/10.1007/s10803-022-05780-z

Yang, T. X., Jia, L., Zheng, Q., Allen, R. J., & Ye, Z. (2018). Forward and backward recall of serial actions: Exploring the temporal dynamics of working memory for instruction. *Memory & Cognition, 47*(2), 279−291. https://doi.org/10.3758/s13421-018-0865-x

Yu, Z. Y. (2014). *The influence of physical features on the subject performed task advantage effects in action memory*. [Unpublished master's thesis]. Jilin University).

[于战宇. (2014). *物理属性在动作记忆 SPT 优势效应中的作用* (硕士学位论文). 吉林大学.]

Modality dependence in the subject-performed task effect for verbal instructions: The moderating role of action−object pair conventionality

XIE Tingting[1], YANG Lianyu[1], WU Xiaotian[1], ZHOU Diandian[1], YU Zhanyu[2]

([1] *School of Psychology, Fujian Normal University, Fuzhou* 350117, *China*)

([2] *School of Education Science, Jiangsu Normal University, Xuzhou* 221116, *China*)

Abstract

The subject-performed task (SPT) effect refers to the finding that action encoding (physically performing described actions) yields better recall of verbal instructions than verbal encoding (merely listening to them). Multimodal theory proposes that this advantage arises because action encoding recruits additional sensory channels beyond the verbal modality, including kinesthetic (motor execution), visual (e.g., observing objects and actions), and tactile input (e.g., touching object surfaces), thereby enriching memory traces and enhancing their retrievability.

Building on this framework, the present study examined how the conventionality of action−object pairings in verbal instructions moderates the multimodal dependencies of the SPT effect, with the aim of distinguishing critical from auxiliary sensory contributions. Three experiments were conducted using familiar everyday objects.

In Experiment 1 (N = 200), we manipulated instruction conventionality (conventional, e.g., "sharpen a pencil"; non-conventional, e.g., "poke a pencil") and the presence versus absence of object pictures (providing visual cues without physical interaction). Results showed that conventional instructions elicited a SPT effect (i.e., performance in the action encoding condition was superior to that in the verbal encoding condition) regardless of whether pictures were present. In contrast, non-conventional instructions failed to produce the SPT effect even when object pictures were available. We hypothesized that static visual input alone may be insufficient to support the online integration of verbs and nouns for non-conventional pairings, and that physical interaction with objects might be necessary. To test this hypothesis, Experiment 2 (N = 58) replaced pictures with real, manipulable objects, enabling participants to physically interact with them during encoding. Under these conditions, non-conventional instructions yielded a significant SPT effect when objects were present but not when absent, indicating that sensorimotor interaction, rather than mere visibility, is essential for the SPT effect with non-conventional instructions. However, because real objects concurrently engage both vision and touch, Experiment 2 could not isolate which sensory channel drives this effect. Experiment 3 (N = 54) addressed

this by orthogonally manipulating visual access (blindfolded vs. unblindfolded) and fine tactile fidelity (wearing gloves vs. bare-handed) during object interaction, with a verbal-only baseline. Results showed that the SPT effect emerged only when vision was available, irrespective of fine tactile condition. Specifically, recall performance was significantly better in both “vision + touch” and “vision + no-fine-touch” action conditions than in the verbal baseline (ps < .001), whereas “no-vision + touch” and “no-vision + no-fine-touch” conditions did not differ from verbal encoding (ps > .999). These findings demonstrate that vision, rather than fine tactile input, is the necessary channel for the SPT effect with non-conventional instructions.

Collectively, these findings reveal that the sensory channel dependence of the SPT effect varies with instruction conventionality. For conventional instructions, kinesthetic input alone, without visual or tactile support, suffices to drive the SPT effect. For non-conventional instructions, the SPT effect requires the coordination of kinesthetic and visual input, while fine tactile information plays no independent role. These results refine multimodal theory by demonstrating that channel contributions are not simply additive, where more channels yield greater benefit, but are instead flexibly determined by whether pre-existing action−object associations are available. This advances the understanding of multimodal memory and offers practical guidance for memory support strategies in populations with sensory or neurodevelopmental differences (e.g., autism spectrum disorder), where optimizing visual−motor integration may enhance instruction memory without relying on tactile precision.

Keywords　SPT effect, instruction memory, multimodal processing theory