

题目：基于价值的议程对学习时间分配影响的眼动研究

作者：姜英杰，王志伟，郑明玲，金雪莲

---

### 第一轮

审稿人 1 意见：

意见 1：在“实验 1 习惯性反应对学习时间分配的影响”中，“习惯性反应”应该是自变量。那么，在实验中是如何测量和控制的呢？实验将词对位置(左、右)作为自变量，这如何反映出“阅读习惯”？从实验结果来看，左侧优势比较明显，这是典型的注意左侧优势效应，应该和“阅读习惯”关系不大。此外，眼睛的注视位置在很大程度上就是注意的位置(特定条件下的 covert attention 除外)。因此，建议整个研究可以从“注意偏向”角度讨论更合适，“阅读习惯”过于笼统。

回应：

学习时间分配 (study time allocation) 领域的研究发现，当同时呈现多个项目时，英语被试(具有从左到右的阅读习惯)优先学习左侧项目，阿拉伯语被试(具有从右到左的阅读习惯)优先学习右侧项目(Bjork, Dunlosky&Kornell, 2013; Dunlosky & Ariel 2011a; Dunlosky & Ariel 2011b; Dunlosky, & Thiede, 2004)，研究者将这一现象定义为习惯性反应(habitual responding)。并发现学习者会根据自身阅读习惯来进行学习时间分配，它通常由词对位置线索(左、右)激发(Dunlosky & Thiede, 2004)。这种习惯性反应与评审专家所提到的注意左侧优势效应实质上是一致的，只不过在本领域的研究中通常以“习惯性反应”来进行表述。

因此，“实验 1 习惯性反应对学习时间分配的影响”旨在通过词对位置(左、右)的设定，来考察具有从左到右阅读习惯的以汉语为母语的中国大陆被试，是否会因为习惯性反应的影响，在词对价值相等的条件下，优先学习左侧项目并给左侧项目分配更多的学习时间。以此作为实验 2 (小分值梯度)和实验 3 (大分值梯度)的基线条件。所以自变量为词对位置(左、右)，因变量为不同位置词对的记忆成绩和学习时间分配的各项眼动指标(优先学习项目、学习时间)。如果被试对左侧位置词对的记忆成绩和各项学习时间分配的眼动指标都要显著高于右侧项目，则说明在项目价值相等的条件下，习惯性反应影响学习时间分配。如果左侧项目的各项因变量指标并不高于右侧项目，则说明项目价值相等的情况下，被试的学习时间分配也不受习惯性反应的影响。从本研究实验 1 的结果来看，项目价值相等的条件下，习惯性反应影响着学习时间分配(左侧项目的记忆成绩和学习时间分配的各项眼动指标均显著高于右侧项目)。

意见 2：词的大小不用像素，而是用视角(度)表示。

回应：对刺激大小的描述可以采用视角也可以采用像素(Tullis, Benjamin, & Liu, 2014)或者尺寸(Castro & Wasserman, 2013)，在提供了屏幕分辨率或屏幕尺寸以及被试距离屏幕距离的条件下，三者可以相互换算。但是考虑到审稿专家一定是站在心理学报对刺激大小描述的表达习惯上建议改成视角的，所以采纳专家意见，将像素改为了视角：“词对呈现视角为 10.40 度×2.17 度”。

意见 3：在“优先选择项目”分析中，选择前半注视点作为分析对象的理由似乎不是很充分。它的作用也不是很大。重点应该是分析“不同阶段”眼动特点。

回应：

(1) 关于“优先选择项目”中前一半注视点的选择理由。

第二位评审专家也关注了同一问题,并建议用中位数而不是平均数来划定前一半注视点。此处按照第二位专家观点对前一半注视点的选取方法做了修改,以中位数为依据来选取前一半注视点。虽然从数据分布来看,并未有极端值出现,但是采用中位数进行分半,确实比平均数更能防范这种风险。

具体处理步骤:(1)记录被试每个试次的注视点个数;(2)计算每个被试 12 个试次注视点个数的中位数;(3)在每个试次中选前“中位数/2”个注视点,作为每个试次的前一半注视点。如,一个被试在某一实验条件下 12 个试次注视点个数的中位数是 10,则取每个试次前  $10/2=5$  个注视点,作为前一半注视点;(4)对前一半注视点落在不同位置(或分值)的项目上的比例进行配对样本(同一实验中不同位置的比较)或独立样本(实验间同一位置比较) $t$ 检验,以考察在不同自变量水平上,优先选择项目是否存在差异。

(2) 关于分析前一半注视点的作用。

“优先选择项目”是学习时间分配研究中的重要指标 (Son & Sethi, 2006),旨在考察自定步调学习中,被试在学习的初始阶段将心理资源优先分配到何种项目上,是认知适应性的反映。本研究通过对前一半注视点的分析,可以考察在等价值、小分值梯度和大分值梯度条件下,被试优先选择的项目是否符合高效学习及认知适应的要求:即分值相等时,按照习惯性反应优先选择左侧项目;而当分值不同时,尤其是高分值项目出现的位置和习惯性反应优先选择的位置不同时,被试是否可以做出符合认知适应的选择,优先选择高分值项目进行学习,克服习惯性反应的影响,提高学习效率。因此,对于前一半注视点的分析,在本研究中非常重要,不能不做。

而且这种分析同专家建议的“重点分析不同阶段眼动特点”并不矛盾,只是作用不同。本研究也对不同阶段各项目的眼动数据进行了分析(见文稿 2.3.2、3.3.2、4.3.2),其目的是考察不同位置项目上不同价值的设定对学习时间分配的不同影响,发现学习时间分配具有情境特异性:在左右同等分值条件下,学习者经过一个短暂的项目选择阶段(前 0.3s)后,迅速进入定向学习阶段,集中注意资源学习左侧项目。在 1 分-5 分条件下,学习者需要经过一个较长的项目选择阶段(前 1.8s),才进入定向学习阶段,对右侧 5 分项目进行学习。而在 5 分-1 分和大分值梯度(1 分-10 分和 10 分-1 分)下,学习者没有经过项目选择阶段,直接进入定向学习阶段,对大分值项目进行学习。此外,定向学习阶段的时长也受到情境的影响。在左右同等分值条件下,定向学习阶段时长约 1.5s;在 1 分-5 分条件下,定向学习阶段只有 1.2s;而在 5 分-1 分和大分值梯度(1 分-10 分和 10 分-1 分)下定向学习阶段明显更长,达到了 1.8s 和 2.1s。该结果表明随着学习情境的改变,基于议程的学习时间分配过程也会发生相应的改变。

**意见 4:** 在实验中,每个词对的学习时间是 3s 钟,时间是限定的。因此,这里并不是“自定步调学习”。建议仅保留表 2 或图 1 中的一个,并将阶段 1、2、3……改成时间(ms)。

**回应:**

(1) 关于本研究是否是自定步调学习

自定步调学习 (self-paced learning) 是相对于固定步调学习 (fixed-paced learning) 而言的。在固定步调学习中,每一项目的学习时间是主试设定的;而自定步调学习中,每一项目的学习时间可以完全或者一定程度上由被试自己决定。本研究中每屏同时呈现 2 个词对,共 3s,学习的总时间由主试设定,但被试可以决定将 3s 的总时间如何在每个词对上进行分配。这是一种自定步调学习的经典研究范式 (Metcalf, 2002)。

(2) 关于图表保留问题。

采纳专家建议,去除“各学习阶段中不同项目的注视时间比例”的描述统计表。

(3) 关于图中横坐标的单位问题。

采纳专家意见,将折线图的横坐标由“阶段”改为“时间”。但单位没有采用 ms,而是采用 s。因为如改成 ms,折线图的横坐标轴刻度标签拥挤,不易清晰标出。

意见 5: 记忆成绩缺少描述统计值。

回应: 文中记忆成绩的描述统计值为项目“正确回忆个数”, 所以实验 1、2、3 记忆成绩描述统计值见表 1、表 2、表 3 中的“正确回忆个数”。

意见 6: 在实验 2 和实验 3 中, 如何定义词的分值?

回应: 实验 2、3 中项目分值的定义参考了先前研究 (Ariel & Dunlosky, 2013; 牛勇, 邱香, & 傅小兰, 2010)。为了便于把本研究的结果与前人研究结果进行比较, 本研究分值梯度分别选取了小分值梯度的“1 分、5 分”和大分值梯度的“1 分、10 分”。并将分值标记在词对上方 (见下图), 通过指导语要求被试尽可能获得高的总分。



审稿人 2 意见:

意见 1: 前言部分可以再组织一下, 让研究背景和目的更加清晰。另外, 第二段的 4-7 行有错误。

回应: 感谢专家评审建议。已对前言文字进行了重新组织, 详见文中蓝色文字。原文第二段 4-7 行的错误也已修正。

意见 2: 实验方法: 实验 2、3 的“一致”和“不一致”条件是 blocked 还是 mixed? 作者需要报告。

回应: 是 block 间的。接受评审专家意见, 已在文中“3.2.4 程序”部分增加了说明。

意见 3: 所有实验的“优先选择项目”分析: 作者根据“平均”注视点个数选择“前半”注视点来分析被试优先选择那些项目来加工。这里“前半”应该根据中位数选择(median-split), 而不是平均数。

回应: 赞同专家观点, 以中位数为依据来选取前半注视点。虽然从数据分布来看, 并未有极端值出现, 但是采用中位数进行分半, 确实比平均数更能防范这种风险。

具体处理步骤: (1)记录被试每个试次的注视点个数; (2)计算每个被试 12 个试次注视点个数的中位数; (3)在每个试次中选前“中位数/2”个注视点, 作为每个试次的前一半注视点。如, 一个被试在某一实验条件下 12 个试次注视点个数的中位数是 10, 则取每个试次前  $10/2=5$  个注视点, 作为前半注视点; (4)对前半注视点落在不同位置 (或分值) 的项目上的比例进行配对样本 (同一实验中不同位置的比较) 或独立样本 (实验间同一位置比较)  $t$  检验, 以考察在不同自变量水平上, 优先选择项目是否存在差异。

意见 4: 正常情况下眼跳每秒发生 2-3 次。在本实验中, 单个试次的学习项目呈现时间仅为 3s, 注视点的数量平均在 12-13 个。作者把学习过程划分为 10 个阶段, 是否偏多?

回应: 本研究将 3s 的学习过程分为 10 个阶段主要是基于下述考虑:

(1) 本研究对学习进程的划分目的是为了对自定步调学习中, 不同条件下两个项目上的学习时间分配情况进行尽量接近时时的探索性监控, 以有利于及时把握动态变化特点和趋势, 从中分析出自定步调学习时间分配的三阶段(项目选择、定向学习和调整阶段)。所以, 将其分成 10 个阶段的这种细致划分, 可以为研究目的达成提供支持。

(2) 10 个阶段的划分并非按照注视点来划分的, 而是分别对每一个 300ms 内各项目的注视时间比例进行的考察。

**意见 5:** 统计方法: 实验 2、3 有两个因素, “优先选择项目”、“总学习时间”和“记忆成绩”的分析需要先做 ANOVA。另外, 三个实验的“各阶段学习时间”分析, 作者对学习的“时程”无很强的事先假设, 应矫正“多重比较”。

**回应:**

(1) 关于没做 ANOVA, 而是采用  $t$  检验的原因。

虽然实验 2 和实验 3 中涉及到项目价值和项目位置, 但该设置是为了形成习惯性反应和议程对学习时间分配影响一致(高分值项目出现在左侧)和不一致(高分值项目出现在右侧)的两个实验条件, 并不是 2(项目位置)  $\times$  2(项目分值) 的两因素实验设计, 如果用 2(项目位置)  $\times$  2(项目分值) 的 ANOVA, 则需要探索性分析大量的事后多重比较结果, 而这些结果中只有一少部分同本研究的假设直接相关。

实验 2 和实验 3 均分别在议程与习惯性反应对学习时间分配影响一致和不一致两种条件下考察议程和习惯性反应共同作用下的学习时间分配特点。主要关注两个问题: ①在议程与习惯性反应对学习时间分配影响一致和不一致两种条件下, 分值梯度能否激发议程。②不一致条件下分值梯度激发的议程能否克服习惯性反应的影响, 使学习者建立起优先学习高价值项目的议程。这两个问题, 均可通过  $t$  检验结果来实现。

问题①, 可用一致(不一致)条件下左侧(右侧)高分项目的“优先选择比例”、“总学习时间”和“记忆成绩”与基线条件(实验 1)左侧项目进行独立样本  $t$  检验。如果一致(不一致)条件下左侧(右侧)高分项目的“优先选择比例”、“总学习时间”和“记忆成绩”与显著高于基线条件(实验 1)左侧项目, 则表明分值梯度激发了议程。

问题②, 可对不一致条件下左侧低分和右侧高分项目“优先选择比例”、“总学习时间”和“记忆成绩”进行配对样本  $t$  检验。如果左侧低分项目“优先选择比例”、“总学习时间”和“记忆成绩”显著高于右侧高分项目, 则表明分值梯度没有能够克服习惯性反应对学习时间分配的影响; 如果左侧低分项目的“优先选择比例”、“总学习时间”和“记忆成绩”与右侧高分项目无显著差异, 则表明分值梯度克服了习惯性反应的影响, 但没有使学习者建立起优先学习高价值项目的议程; 如果右侧高分项目“优先选择比例”、“总学习时间”和“记忆成绩”显著高于左侧低分项目, 则表明分值梯度是否克服了习惯性反应的影响, 并使学习者建立起优先学习高价值项目的议程。

综上, 本研究具有明确的实验假设, 用  $t$  检验就能更为直接地验证相关假设, 解决本研究关注的问题。

(2) 关于三个实验“各阶段学习时间”分析不进行“多重比较”的原因。

在“各阶段学习时间”分析中, 研究关注的问题是“学习者在各个阶段更多的将注意资源分配给何种特征的项目”。因此, 相对于对某类项目注视时间进行 10 个阶段的多重比较, 分别对 10 个阶段不同项目间的注视时间进行差异检验更有利于说明问题。当不同项目注视时间出现显著差异时, 提示我们自定步调学习进入了下一个阶段, 这是对不同条件下自定步调学习过程进行阶段划分(项目选择阶段、定向学习阶段和调整阶段)的重要依据。此外, 对某类项目注视时间进行 10 个阶段的多重比较会出现  $C_{10}^2=45$  种比较结果, 对结果描述困难,

不利于读者理解。

**意见 6:** 实验 2、3 的结论: 被试选择左侧项目的偏向在小分值“不一致”条件下“消失”(实验 2), 在分值“不一致”条件下“被反转”(实验 3)。实验 2 的结论“分值梯度...没有能够克服习惯性反应的影响”有误。

**回应:** 接受审稿专家意见, 研究结论修改为“(2)小分值梯度能够克服习惯性反应的影响, 但不能促使被试建立起优先学习高价值项目的学习议程; 大分值梯度能够克服习惯性反应影响, 且能够促使被试建立起优先学习高价值项目的学习议程。”

**意见 7:** 实验 3 的结论“相较于小分值梯度, 大分值梯度下学习者项目选择决策的时间更短”。该结论需要有统计结果支持。具体来讲, 实验(2 vs. 3) × 分值梯度(小 vs. 大) × 学习阶段(1-10)的 3-way 交互作用显著, 作者才能做出该结论。

**回应:** 接受审稿专家的指正。

将原文中“相较于小分值梯度, 大分值梯度下学习者项目选择决策的时间更短”的结论修改为, “不一致条件下, 随着分值梯度变大, 学习者项目选择决策的时间有缩短的趋势 (1 分-5 分条件下 1.8s, 1 分-10 分条件下 0s)。”

专家建议考察“实验(2 vs. 3) × 分值梯度(小 vs. 大) × 学习阶段(1-10)的 3-way 的交互作用”, 实际上, 实验 2 为小分值梯度条件, 实验 3 为大分值梯度条件, 所以专家的建议其实就是考察分值梯度(小 vs. 大) × 学习阶段(1-10)的 2-way 的交互作用。但是每个阶段的学习时间均为 3s 的十分之一 (0.3s), 所以按照 10 个阶段分析无法得出何种分值梯度下项目选择阶段更短结论。

因此专家修改意见的目的是为“考察小分值梯度 (实验 2) 和大分值梯度 (实验 3) 条件下项目选择阶段的时长是否存在显著差异”。若进行差异检验, 分值梯度为自变量, 项目选择阶段时长为因变量。本研究项目选择阶段的界定方法为: 将优先项目(实验 1 中指左侧项目; 实验 2、3 中指高分值项目)学习时间比例首次显著高于非优先项目的阶段定义为定向学习阶段的开始; 将定向阶段之前各项目学习时间无显著差异的阶段定义为项目选择阶段。具体步骤如下: (1) 将每种实验条件下每屏总学习时间(3s)平均分为 10 个阶段, 分别计算各阶段中左侧和右侧项目的注视时间比例。(2) 分别对左侧和右侧项目 10 个阶段的注视时间比例进行配对样本  $t$  检验。(3) 将优先项目学习时间比例首次显著高于非优先项目阶段之前的阶段定义为项目选择阶段。例如, 实验 1 中, 阶段 1 左侧项目和右侧项目的学习时间比例无差异, 阶段 2 中左侧项目学习时间比例显著高于右侧项目,  $t_{\text{阶段 2}}(14) = 4.23, p < 0.001, \text{Cohen's } d = 1.09$ , 项目选择阶段时长即为阶段 1 的时长 300ms。数据分析发现, 小分值梯度下, 一致条件 (5 分-1 分) 和不一致条件 (1 分-5 分) 下项目选择阶段时长分别为 0s 和 1.8s; 大分值梯度下, 一致条件 (10 分-1 分) 和不一致条件 (1 分-10 分) 下项目选择阶段时长都为 0s。每种实验条件下的项目选择阶段只有一个数值, 所以无法进行统计学上的差异检验, 确实无法得出“相较于小分值梯度, 大分值梯度下学习者项目选择决策的时间更短”的结论。

但是, 在不一致条件下, 高分值梯度的项目选择阶段时长 (0s) 在数值上明显短于小分值梯度的项目选择阶段时长 (1.8s), 所以将原文中“相较于小分值梯度, 大分值梯度下学习者项目选择决策的时间更短”的结论修改为, “不一致条件下, 随着分值梯度变大, 学习者项目选择决策的时间有缩短的趋势 (1 分-5 分条件 1.8s, 1 分-10 分条件下 0s)。”

**意见 8:** 英文文题和摘要还有很大提升的空间, 审稿人的部分修改建议见附件。

**回应:** 接受审稿专家的修改意见。

意见 9: 请注意“全角”和“半角”括号的差别。

回应: 感谢专家提醒, 已参照《心理学报》的写作要求将文中的全角括号改为了半角括号。

意见 10: 数学符号(=, >, <)前后要有空格。

回应: 接受专家指导, 在数学符号(=, >, <)前后增加了空格。

---

## 第二轮

审稿人 1 意见:

意见 1: 关于“自定步调学习”。作者将实验的学习阶段的两个词看成是 2 个项目, 而在线索回忆阶段, 却将两个词看成是 1 个项目, 因为只有在学习阶段将两个词之间建立联系, 才能写出与线索词对应的目标词。Metcalf (2002) 并未提及“self-paced learning”。

回应:

(1) 学习阶段每个“线索词-目标词”词对是一个项目, 每个词对由两个词组成。

(2) Metcalf (2002) 的文章中, 确实没有直接提及“self-paced learning”。综合考虑第一和第三位审稿人对该问题的意见, 作者在文中去掉了“自定步调”的表述。在此对审稿人的细致和耐心表示感谢!

意见 2: 图 1 中的横坐标没有单位。“阶段”也与坐标轴的属性不一致。相应的统计结果描述与图 1 不能对应, 图中横坐标用时间, 而文字描述中仍然用 1、2、3 等阶段。图 2 等也存在同样问题。

回应:

(1) 感谢审稿人的建议, 图 1、图 2、图 3、图 4、图 5 的横坐标均已标注单位, 并将横坐标轴名称改为“时间”。

(2) 所有与图相应的统计结果描述均已按照审稿人建议, 将“阶段 n”修改为了“m~ns”。

意见 3: “学习时间”本身过于宽泛。从习惯反映讨论学习时间分配问题也过于笼统。并未能深入到具体的心理过程。因此, 文章无法深入描述学习时间分配的心理机制。眼睛注视点分布在很大程度上反应了注意的分配特性。在上次意见中, 曾建议文章从注意角度讨论。

回应: 学习过程是一个综合的认知加工过程, 需要注意资源参与, 作者赞同审稿专家“眼睛注视点分布在很大程度上反应了注意的分配特性”的观点, 也正是利用这一点, 通过注视点的分布来考察学习时间分配情况的。接受审稿专家建议, 在此次修改过程中, 全文增加了从注意角度的论证。

审稿人 2 意见:

意见 1: 在第一轮审稿意见中, 我建议作者在进行  $t$  检验之前先做 ANOVA。作者在回复中用了大段的篇幅来解释为何自己没有做 ANOVA。这些回复并没有回答审稿人提出的问题。实际上, 作者给出的解释恰恰说明本研究在因变量的选择方面存在问题。作者在图 1-5 中比较的是注视“左”和“右”侧的百分比, 这等于有两个因变量。实际上“左”= 1-“右”, 更有意义的方法是给出一个“左侧偏向”的度量, 即“左-右”,  $> 0$  表明有左侧倾向,  $= 0$  表明没有左侧倾向,  $< 0$  表明有右侧倾向。这样就可以比较高额奖励与“左侧”一致或不一致条件下的“左侧偏向”, 也就是直接比较图 2 和 3, 4 和 5。

回应: 感谢并接受审稿人建议。统计了“左-右”的“左侧偏量”, 并对不同实验条件间的“左

一右”的“左侧偏量”进行了比较。

**意见 2:** 作者在回复中讲修改了英文摘要，但是审稿人未见有变化。

**回应:** 由于作者的疏漏，第二稿中确实没有将修改完的英文摘要附上，第三稿中已对此进行了改正。

**审稿人 3 意见:**

**意见 1:** 作者对审稿专家的意见进行了很好的回答。具体情况如下：对第一位审稿专家的建议回答情况是这样的：

对第 1 个问题的回答，可以接受。实际上就是没有控制和测量习惯性反应。已经认为中国被试的习惯性反应就是从左向右的阅读习惯。

**意见 2:** 对第 2 个问题的回答，用不着固执己见，既然接受专家意见，修改就是了。可以接受。

**回应:** 已在文中按照审稿人意见进行了修改。

**意见 3:** 对第 3 个问题的回答，能结合第二个审稿人意见进行修改，很好。说明作者学术严谨性很高。

**意见 4:** 对第 4 个问题的回答，“自定步调学习”的说法，可以修改一下，因为在眼动研究中，有自己的理解。我们普遍把时间由被试自己控制的，理解成“自步速”。其他的修改接受。

**回应:** 综合考虑第一位审稿人和第三位审稿人对该问题的意见，接受两位审稿人意见，在文中去掉了“自定步调”的表述。

**意见 5:** 第 5 题的回答，可以接受。需要在论文中写清楚。不要让读者猜。

**回应:** 已经文中写清楚。

**意见 6:** 第 6 题的回答可以接受。需要在论文中写清楚。

**回应:** 已经文中写清楚。

**意见 7:** 对第二位审稿专家建议的回答情况如下：对第 1 题的回答，已经修改，接受。

**意见 8:** 对第 2 题的回答，已经修改，接受。

**意见 9:** 对第 3 题的回答，同意。

**意见 10:** 对第 4 题的回答，清楚。可以接受。

**意见 11:** 对第 5 题的回答，清楚，可以接受。以后研究中最后不要有 10 个取值，确实太多。

**意见 12:** 对第 6 题的回答，作者修改了文中的错误，文章的科学性提高。接受。

意见 13: 对第 7 题的回答, 作者修改了文中的错误, 科学性提高。其他解释可以接受。后面的问题, 也进行了修改, 接受。

### 第三轮

#### 审稿人 1 意见:

意见 1: 实验 1: 主要目标是建立“是否有左侧优先加工的偏向”。第一个分析 (2.3.1) 可以用简单  $t$  检验 (左 vs. 右), 说明被试在表 1 所列出的三个因变量方面都有差异 (左>右)。既然有了这个明显、一致的左侧偏向, 我们就可以把这个偏向数量化 ( $d_{左-右}$ ), 然后分析在没有奖赏的正常情况下, 这个偏向是如何随学习时间变化的。

因为涉及到时程, 不能用每个时间点的  $t$  检验结果来说明, 必须使用单因素 (时间) ANOVA 对  $d_{左-右}$  进行分析。只有时间的主效应显著, 我们才能得出“左侧偏向”确实随时间有变化。然后才可以呈现现在文中报告的事后  $t$  检验, 来确定偏向出现在学习初期, 中段, 还是晚期。【10 个阶段两两比较】图 1 需要修改, Y 坐标应该为  $d_{左-右}$ , 把 Y=0 标出来, 表示没有偏向, Y+ 标为左偏向, Y- 标为右偏向。

#### 回应:

(1) 接受审稿专家的建议, 用配对样本  $t$  检验 (左 vs. 右) 对优先选择项目 (见文中 2.3.1) 和记忆成绩 (见文中 2.3.3) 进行了考察, 发现在等分值条件下, 被试在上述因变量指标上左侧项目均优于右侧项目, 且差异显著。上述结果验证了实验 1 的预期, 左侧优先加工偏向存在。

(2) 接受审稿专家的建议, 为进一步考察左侧偏向是否随时程变化, 以时间段为自变量, 对左、右侧项目注视时间比例之差  $d_{左-右}$  进行单因素重复测量方差分析 (见 2.3.2)。结果发现, 时间段主效应显著, 表明左侧偏向随学习时程而变化, 在等价值条件下, 不同位置项目学习时间分配是一个动态变化的过程。之后进行了事后检验 (LSD) (见下表)。

表 1 同等分值条件下, 多重比较(LSD)结果

时间段	0.0~0.3s	0.3~0.6s	0.6~0.9s	0.9~1.2s	1.2~1.5s	1.5~1.8s	1.8~2.1s	2.1~2.4s	2.4~2.7s	2.7~3.0s
<b>0.0~0.3s</b>										
<b>0.3~0.6s</b>	-.71***									
<b>0.6~0.9s</b>	-1.03***	-.32**								
<b>0.9~1.2s</b>	-1.11***	-.39*	-0.08							
<b>1.2~1.5s</b>	-1.09***	-.38*	-0.06	0.01						
<b>1.5~1.8s</b>	-.63*	0.09	0.40	.48*	.47*					
<b>1.8~2.1s</b>	-0.18	0.53	.84**	.92**	.91**	.44**				
<b>2.1~2.4s</b>	-1.03***	-.32**	0.00	0.08	0.06	-0.40	-.85**			
<b>2.4~2.7s</b>	0.27	.99**	1.30***	1.38***	1.37***	.90**	0.46	1.30***		
<b>2.7~3.0s</b>	-1.45***	-.74***	-.42**	-0.34	-.36*	-.82**	-1.27**	-.42**	-1.72***	

结果表明, 大部分两两时间段的  $d_{左-右}$  存在显著差异, 进一步说明左侧偏向程度随学习进程而变化, 学习时间分配是一个动态变化的过程。这和时间段主效应显著所得结论一致。由于篇幅限制, 所以文中未呈现事后检验的结果。

本研究除了上述目的, 还要对动态变化中的学习时间分配进行阶段划分, 考察不同情境下, 学习时间分配是否存在项目选择、定向学习和学习调整阶段。具体到在实验 1 中, 首次出现左侧偏向的时间段定义为定向学习阶段的开始, 左侧偏向首次消失的时间段定义为学习调整阶段得开始, 将首次出现左侧偏向前所有时间段定义为项目选择阶段。这种阶段划分,



需要数据支持，而上述方差分析及事后检验结果虽然可以比较两两时间段  $d_{左-右}$  的差异，但无法确定每个时间段内是否出现了左侧偏向。例如时间段 0.0~0.3s 和时段 0.3~0.6s 之间的  $d_{左-右}$  差异显著，只表明两个时间段左侧偏向的指标  $d_{左-右}$  存在差异，但并不能说明哪个时间段出现了左侧偏向。为了找到首次出现左侧偏向的时间段，需要用单样本  $t$  检验考察每个时间段的  $d_{左-右}$  与 0 是否存在差异，并配合文中图 1，将优先(实验 1 的左侧项目)与非优先项目学习时间比例之差首次显著大于 0 的阶段(左侧偏向出现)定义为定向学习阶段的开始；将优先与非优先项目学习时间比例之差首次不显著大于 0 的阶段(左侧偏向消失)定义为调整阶段的开始；将定向阶段前  $d_{左-右}$  与 0 无显著差异的阶段定义为项目选择阶段。所以文中保留了每个时间段  $d_{左-右}$  与 0 进行差异检验的分析。由于篇幅限制，没有报出所有结果，仅选择阶段划分的关键节点数据进行了报告(见文中 2.4.2)。实验 2 和实验 3 的阶段划分也是同样处理的(见文中 3.4.2、4.4.2)。

(3) 接受审稿专家的建议，对图 1 进行了修改。纵坐标为  $d_{左-右}$ ，并标出了纵坐标为 0 的位置。

## 意见 2:

实验 2 是考察小梯度奖赏的作用。作者一再强调实验 1 是基线，这个说法有一定的误导性。如果真正需要基线的话，实验 2 应该包含一个无奖赏条件(与实验 1 同)，不然有两个水平是被试内的，一个水平是被试间的，统计上会比较麻烦；当然也可以做两个 ANOVA 分别比较实验 2 的两个条件和实验 1。但是，这个问题不会损害实验的主要结论。

实验 2 的问题与实验 1 相同，作者应该做 ANOVA，否则无法得出两个条件的“时程”有差别的结论。这是个简单的被试内两因素 ANOVA，2(高奖赏在左 vs. 高奖赏在右) X 10(时间段)，因变量同样是  $d_{左-右}$ 。从图 2、3 来看，2-way 交互应该显著。然后作者应该用  $t$  检验比较两个条件的  $d_{左-右}$ ，而不是分别检验两个条件下各个时间点的  $d_{左-右}$  是否与 0 有差异。图 2 和图 3 应该合并 1，显示  $d_{左-右}$  在两个条件下如何随时间变化。

## 回应:

(1) 接受审稿专家的意见，为了避免误导，在正文中去掉了实验 1 作为基线条件的表述。实验 1 实际上是一种等分值条件，学习时间分配只受被试在长期学习过程中形成的习惯性位置偏好的影响，在实验 1 中表现为受项目位置的影响。

(2) 接受审稿专家建议，用 ANOVA 分别对实验 2 小分值梯度下的两个条件和实验 1 比较，以考察分值激发的议程能否克服位置偏好带来的习惯性反应对学习时间分配的影响，详见文中 3.3。

参照专家上述建议，对实验 3 进行了相同的处理，详见文中 4.3。

(3) 接受审稿专家的建议，对各时间段内左侧项目与右侧项目的注视时间比例之差  $d_{左-右}$  进行 2(大分值位置: 5分-1分; 1分-5分) X 10(时间段)重复测量的方差分析。结果发现，大分值位置与时间段的交互作用显著。然后依照审稿专家建议进行了简单效应分析，用  $t$  检验分别比较了 10 个阶段两个条件的  $d_{左-右}$ ，详见正文 3.3.2 中蓝色字体部分。参照专家建议，对实验 3 也进行了同样处理，详见文中 4.3.2。

同实验 1 一样，因考虑到各时间段  $d_{左-右}$  与 0 进行差异检验是确定每个时间段是否出现左侧偏向，并据此考察学习时间分配过程是否可以划分为项目选择、定向学习和学习调整三个阶段的依据，所以在文中进行了保留，详见正文 3.4.2 中蓝色字体部分。实验 3 中的阶段划分也进行了同样处理(见文中 4.4.2)。

(4) 接受审稿专家的建议，对图 2 和图 3 进行了合并(见文中图 2)。

意见 3: 实验 3 同理。图 4 和 5 也需要合并来比较两种条件的时称差异，Y 坐标为  $d_{左-右}$ 。

回应: 接受审稿专家的意见，已按照专家建议对图 4 和 5 进行合并(见文中图 3)。再次感谢审稿专家对本文提出的建设性修改建议!

## 第四轮

审稿人 1 意见:

意见 1: 中英文文题不对应, 建议更改。

回应: 接受审稿专家建议将中英文标题改为一致:

中文题目: 基于价值的议程对学习时间分配影响的眼动研究

英文题目: How Value-based Agendas affect Study Time Allocation: An Eye Tracking Study

意见 2: 比较图 1-3, 不难看出图 1 中的数据有明显错误。建议作者检查。另外, 图中请添加误差线(error bars, 可以是 SEM 或者 95%置信区间)。

回应:

(1) 经核查数据, 发现图 1 中数据确实有误, 已在文中更正(见图 1)。

(2) 接受审稿专家意见, 图 1-3 均增加了误差线(SEM), 如下图。

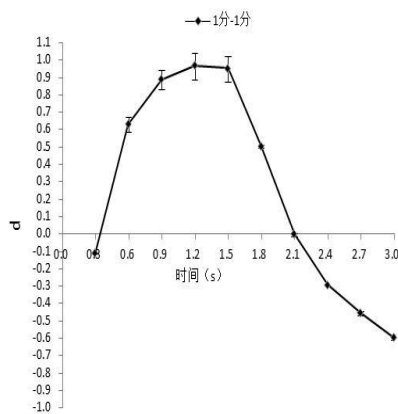


图 1 等分值条件 d 随时程的变化

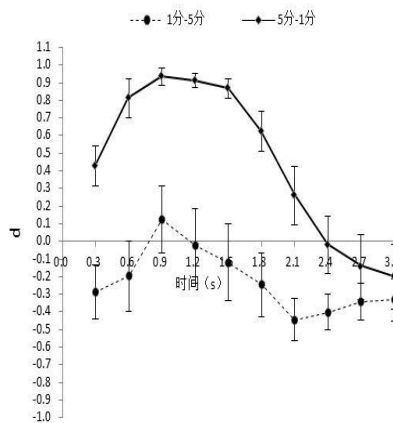


图 2 小分值梯度 d 随时程的变化

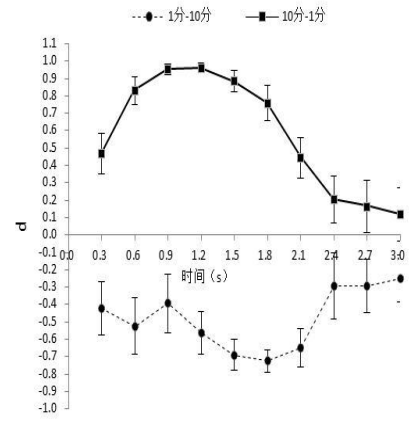


图 3 大分值梯度 d 随时程的变化

注: (1)  $d$  大于 0 表明左侧项目注视时间比例大于右侧, 小于 0 表明右侧项目注视时间比例大于左侧。(2) 误差线为标准误 (SEM)。

意见 3: 统计结果部分的修改建议见审稿稿。

建议删除 3.1.1 中“5-1 条件前半注视点分布在左侧项目的比例显著高于右侧项目,  $t(13) = 19.28, p < 0.001, Cohen's d = 5.15$ , 表明被试优先学习左侧 5 分项目。”因为 5-1 和 1-1 条件都有左侧偏向, “项目位置”的主效应已经说明了这个问题, 无需再检验 5-1 条件。

回应: 接受审稿专家建议, 删掉了对 5-1 条件的检验。

意见 4: 建议删除 3.1.1 中“1-5 条件前半注视点在左、右侧项目上的比例无显著差异,  $t(13) = 0.55, p > 0.05$ , 表明 1-5 条件分值激发的议程虽然克服了习惯性反应的影响, 但还不能达到优先学习高价值项目的程度。”因为 1-5 条件内的比较不能说明奖赏梯度有作用。

回应: 单纯 1-5 条件内的比较确实不能说明奖赏梯度有作用, 但 1-5 条件内的比较目的是为了考察 1-5 条件激发的议程能否克服习惯性反应的影响, 并促使被试优先学习右侧高分项目, 结果表明 1-5 条件激发的议程能够克服习惯性反应的影响, 但议程的作用尚未达到优

先学习右侧高分项目的强度，因此，实验 3 进一步增大分值梯度，考察 1-10 条件能否促使被试建立起优先学习右侧高价值项目的议程。所以保留了该部分的分析。

意见 5: 3.3.2 中建议将“0.0~0.3、0.3~0.6、0.6~0.9、0.9~1.2、1.2~1.5、1.5~1.8、1.8~2.1、2.1~2.4s 时间段，5-1 条件下的  $d_{左-右}$  显著大于 1-5 条件， $t_{0.0-0.3s}(13) = 3.32$ ,  $p < 0.01$ ,  $Cohen's d = 0.89$ ;  $t_{0.3-0.6s}(13) = 4.22$ ,  $p < 0.01$ ,  $Cohen's d = 1.13$ ;  $t_{0.6-0.9s}(13) = 4.43$ ,  $p < 0.001$ ,  $Cohen's d = 1.18$ ;  $t_{0.9-1.2s}(13) = 4.47$ ,  $p < 0.001$ ,  $Cohen's d = 1.19$ ;  $t_{1.2-1.5s}(13) = 4.40$ ,  $p < 0.001$ ,  $Cohen's d = 1.18$ ;  $t_{1.5-1.8s}(13) = 4.79$ ,  $p < 0.001$ ,  $Cohen's d = 1.28$ ;  $t_{1.8-2.1s}(13) = 3.42$ ,  $p < 0.01$ ,  $Cohen's d = 0.91$ ;  $t_{2.1-2.4s}(13) = 2.75$ ,  $p < 0.05$ ,  $Cohen's d = 0.73$ 。”简化为“所有检验:  $t > 2.75$ ,  $p < 0.01$ ,  $Cohen's d > 0.73$ 。”

回应: 接受审稿专家建议，将此处缩写为“在 0.0~2.4s 间的 8 个时间段 5-1 条件的  $d_{左-右}$  显著大于 1-5 条件， $ts \geq 2.75$ ,  $ps < 0.05$ ,  $Cohen's ds \geq 0.73$ 。在 2.4~2.7、2.7~3.0s 时间段，两个条件间  $d_{左-右}$  无显著差异， $ts \leq 1.18$ 。”并同时 3.4.2、4.3.2 和 4.4.2 也进行了类似修改，详见文中蓝色字体。