

# 认知心理学报告



## Stroop 效应实验报告

专业：心理学  
班级：心理 1402 班  
学号：3140104088  
姓名：李文敏  
性别：男

# Stroop 效应实验\*

李文敏<sup>1</sup> 董一胜<sup>\*\*</sup>

(浙江大学心理与行为科学系, 杭州, 310028)

**摘要** 本实验通过中文汉字的词义干扰 Stroop 范式对 52 名浙江大学心理与行为科学系的本科生的 Stroop 效应进行探究, 结果发现: (1) 当颜色和字义两个维度不一致的时候, 会对被试的反应产生干扰, 并且两个实验任务之间的 Stroop 效应没有显著差异 (2) Stroop 不存在性别差异, 也不存在顺序效应, (3) 辨色任务中的 Stroop 效应不存在明显的练习效应, 识字任务中的 Stroop 效应存在一定的练习效应 (4) Stroop 效应受到诸多因素, 例如颜色字义搭配、颜色识别能力、语境因素、年龄因素等。

**关键词** Stroop 效应 性别差异 Stroop 效应的顺序效应 练习效应

## 1. 引言

### 1.1 Stroop 效应的提出

Stroop 效应是由美国心理学家 John Ridley Stroop 于 1935 年发现的。当时, 他在实验中发现, 当命名一个用红墨水写的字的颜色时, 有意义刺激 (如“绿”字) 要比无意义刺激 (如“𠂇”字) 的反应时间更长 (Stroop, 1935)。这种同一刺激的字色信息 (红) 与词义信息 (绿) 相互发生干扰的现象称为 Stroop 效应, 从广义的角度看, Stroop 效应反映的是对一个刺激的两个维度的加工发生相互干扰的现象。

Stroop 是在研究干扰效应 (当时也被称为抑制) 时发现上述现象的。在实验中, 通过比对“字色矛盾”组的字色命名速度与单纯色块组的命名速度, 考察了字义对颜色辨别 (命名) 的影响; 通过对比“字色矛盾”组的彩色字的阅读速度和“字色无关”的黑色字的阅读速度, 考察了字色对字义辨别的影响。进一步的实验则发现, 通过大量练习可以显著降低字义对字色命名的干扰。

### 1.2 Stroop 效应的解释

对于 Stroop 效应的解释, 主要有以下五种理论假设 (MacLeod, 1991)。Stroop 早年的解释接近于早期的相对加工速度和自动化理论。而随着各种理论的发展, 平行分布加工模型是迄今解释 Stroop 效应的最好理论模型。

理论一: 相对加工速度理论 (赛马理论)。这一

理论的依据是字义辨别要快于字色辨别。该理论认为, 人们对刺激的两个维度——字色和字义的加工是平行的, 但加工速度不同, 字义辨别要快于颜色辨别。所以字词的加工先达到反应阶段。如果字词信息与颜色信息一致, 就对颜色辨别产生促进; 相反, 如果不一致, 则对颜色辨别产生干扰。由于颜色辨别晚于字义辨别, 故颜色信息不会对字义辨别产生影响。然而, 该理论不能解释当两刺激维度不同时呈现时所发现的实验结果。

理论二: 自动化理论。该理论区分了自动化加工和控制加工这两个概念。自动化加工是指加工较快, 不需要注意、能随意发生的加工; 而控制加工则较慢, 需要注意的参与和控制。在 Stroop 任务中, 字义加工属于自动加工, 而字色加工则属于控制加工。因此, 字义辨别能对颜色辨别产生干扰而反之则不能。近年来的研究表明, 自动化加工会随学习的进展而呈梯度变化。

理论三: 知觉编码理论。该理论认为, 在 Stroop 效应中, 颜色信息的知觉编码被来自颜色词的不匹配信息所减慢。但是有证据表明 Stroop 效应不仅发生在知觉编码阶段, 而且也发生加工阶段。

理论四: Logan 的平行加工模型。该模型把 Stroop 效应看成是从刺激各维度收集证据进行决策的过程。其中, 刺激的每个维度的加工速度由其权重决定, 而权重又影响每个维度对决策的贡献的大小, 权重越大, 影响也就越大。如果来自某一维度的证据和要求反应维度一致, 就会降低反应阈限, 从

\* 完稿日期: 2016-10-10

\* 研究项目: 认知心理学实验项目

\*\* 通讯作者: 董一胜, 浙江大学心理与行为科学系, E-mail: dongyisheng@zju.edu.cn

而加快该维度的加工速度,反之则会减慢。

理论五:平行分布加工模型(Parallel Distributed Processing, PDP)又称神经网络模型。PDP 系统由很多相互联结的模块组成。每个模块包括许多简单的相互联结的加工单元,每个加工单元负责接收来自其他单元的输入并提供输出。(注意能调节加工单元的各项操作,使其成为另一加工单元的信息源)。每条通路由一组相互联结的模块组成。当 PDP 系统进行任务时,它会选择一条通路,通路中的联结确定了通路的强度,通路的选择从而也确定了信息加工的速度与准确性。PDP 系统的信息加工就是通过激活不同强度的通路传播而进行,由于通路可能重叠,因此,信息加工允许发生干扰或促进现象(交互作用)。

这五种理论尽管只能解释 Stroop 效应中的部分实验,但对 Stroop 效应的解释各有其贡献。在 Stroop 任务中,当命名字的颜色时要求被试忽略字义,但是这个忽略任务总是被显著的加工(对词义加工)所抑制。为了调和这种矛盾,2003 年 Robert 等提出 Stroop 效应的建构理论。建构一词随着建构主义的提出受到许多研究者青睐。在心理学的各领域都存在建构的提法。认知建构、情绪建构、人格建构、语言获得建构等词语频繁出现在相关文献中。我们认为建构一词更强调综合和统一,是一种动态的变化过程。相应地 Stroop 效应的建构理论认为在选择性注意中,基于记忆的两个结构)维度的不平衡性和维度的不确定性引导注意选择那些在刺激维度内或者附近的那些显眼的、奇怪的、或者与任务相关的信息进行加工。维度的不平衡性和维度的不确定性对目标刺激的组合、对分心物抑制的组合进行调节,同时也会对早先刺激的记忆进行调节。建构理论认为在 Stroop 范式中有 4 种不同类型的信息影响 Stroop 效应:呈现刺激的背景、刺激的大小数量、一致性效应以及任务效应。该理论强调:被试对信息的选择所作的努力可能被新信息所中和。如果不相关的信息是显眼的或者是和相关信息是相关的,这样不相关的信息也可能得到加工。作为靶子目标上的字比起在其他地方呈现的字更显眼,因此就有可能调节或改变 Stroop 效应。尽管 Stroop 研究中建构理论是最新的观点,但是研究文献仅有一篇,但有人认为在解释 Stroop 效应的范围和多样性方面被认为优于其他模型。

### 1.3 研究目的

本研究主要是重复 Stroop 等人的经典实验,探

讨 Stroop 效应所产生的可能原因以及内在机制,同时进一步分析影响 Stroop 效应可能的因素。

## 2.方法

### 2.1 方法

被试为浙江大学心理与行为科学系大三本科生 52 名(其中男生 26 名),所有被试的平均年龄为 20 岁(19-21 岁,SD=0.77 岁)。实验被试身体健康,无色弱色盲等视觉障碍。

### 2.2 仪器与材料

IBM-PC 计算机一台,认知心理学教学管理系统。本实验呈现的字符集为“绿”、“红”和“黑”,字符有三种颜色:红色、绿色和黑色,每个字符的大小约为 2.0cm×2.0cm。

### 2.3 实验设计与流程

本实验采用单因素被试内设计。自变量有 3 个水平:字色一致、字色矛盾和字色无关。被试有两个任务:辨色任务和识字任务。辨色任务要求被试对字色做出判断;而识字任务则要求被试对字义做出判断。两个任务的顺序在被试间对抗平衡。单次试验流程见图 11-1。

对于辨色任务:首先在屏幕上中央呈现一个黄色“+”注视点,500~1500 毫秒后在屏幕中央呈现第一个字符,该字符的颜色有可能是红色或绿色,被试的任务是判断该字符是红色还是绿色,并立即做出按键反应。如果是绿色按“J”键;是红色则按“F”键。为了减少被试按键过程中的反应定势,生成的实验序列经 Wald-Wolfowitz 游程检验,显著性大于 0.10(双侧)。

对于识字任务:首先在屏幕上中央呈现一个黄色“+”注视点,500~1500 毫秒后在屏幕中央呈现第一个字符,该字符有可能是“红”字或“绿”字,被试的任务是判断该字符是“红”字还是“绿”字,并立即做出按键反应。如果是“绿”字按“J”键;是“红”字则按“F”键。为了减少被试按键过程中的反应定势,生成的实验序列经 Wald-Wolfowitz 游程检验,显著性大于 0.10(双侧)。

被试做出按键后,会得到相应的反馈,指示被试反应正确与否及反应时。如果被试在字符出现后 1000 毫秒内不予以反应,程序将提示反应超时,告诉被试尽快反应。随机空屏 600~1300 毫秒后,自动进入下一次试验。辨色任务或识字任务实验开始前,从正式实验中随机抽取 20 次作为练习,练习的时候,无论反应正确、错误或超时均有反馈,但结果

不予以记录。练习的正确率达到 90%后方可进入正式实验。正式实验在被试做出正确反应后没有提示，反应错误或反应超时则会有提示。正式实验有 120 次试验，分 4 组（每组 30 次），组与组之间分别

有一段休息时间。正式实验结束后，进入错误补救程序，即将之前做错的试验再次呈现，直到被试全部反应正确为止。整个实验包括辨色任务和识字任务两部分，两者全部完成约 30 分钟。

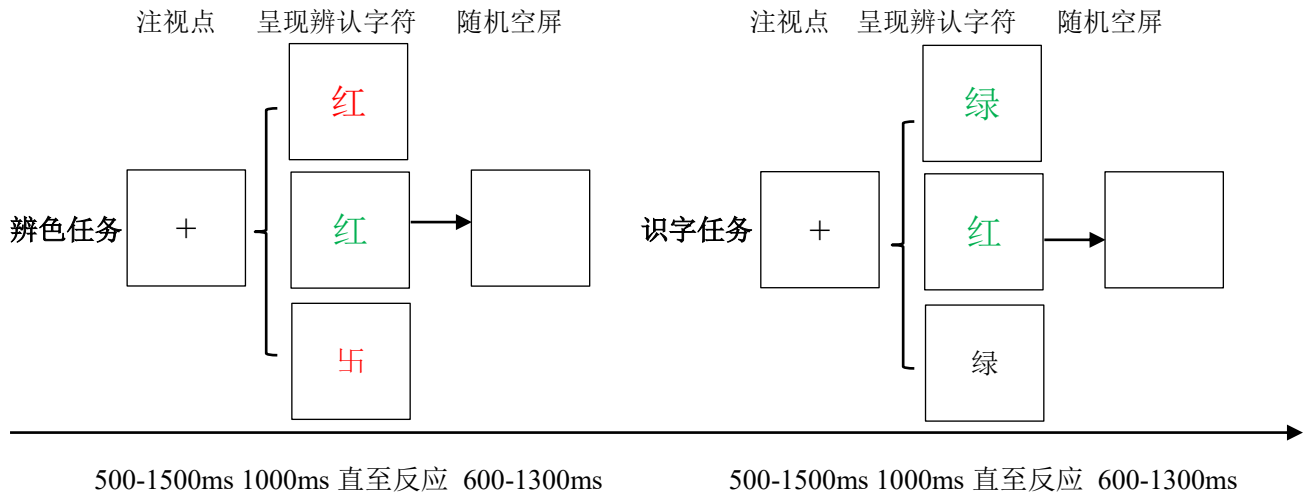


图 2-3-1 Stroop 效应实验流程示意图

### 3.结果

#### 3.1 不同任务下不同字色情况下的平均反应时和错误率

分别计算出每个被试在识别任务和辨色任务下不同字色条件下的平均反应时和平均错误率，详细数据参见附录表 1，所有被试的平均数据见表 3-1。

对实验任务和字色条件的反应时数据进行二因素的重复测量方差分析，球形检验均不符合 ( $p < .01$ )，采用多变量分析结果。结果显示实验任务（辨色任务与识字任务）的主效应显著 ( $F(1,51) = 6.487$ ,  $p < .05$ ,  $\eta_p^2 = 0.113$ )，辨色任务的平均反应时 ( $M = 448.22\text{ms}$ ) 低于识字任务的平均反应时 ( $M = 464.44\text{ms}$ )，但是两者差异并不大，并且效应量只有 0.113，可该效应并不强；字色条件的主效应显著 ( $F(2,50) = 13.536$ ,  $p < .01$ ,  $\eta_p^2 = 0.351$ )，字色矛盾下的平均反应时大于其他两种条件下的平均反应时，具有中等效应；两者的交互作用不显著 ( $F(2,50) = 2.237$ ,  $p > .10$ )，即两个实验的 Stroop 效应差异不显著。

对实验任务和字色条件的错误率数据进行二因素的重复测量方差分析，球形检验均不符合 ( $p < .01$ )，采用多变量分析结果。结果显示实验任务（辨色任务与识字任务）的主效应显著 ( $F(1,51) = 4.651$ ,  $p < .05$ ,  $\eta_p^2 = 0.084$ )，辨色任务的平均错误率 ( $M = 0.05$ ) 低于识字任务的平均反应时 ( $M = 0.063$ )，但是两者差异并不大，并且效应量只有 0.084，可该效应非常微弱，很可能是由于本研究的大样本造成的；字色条件的主效应显著 ( $F(2,50) = 10.193$ ,  $p < .01$ ,  $\eta_p^2 = 0.290$ )，字色矛盾下的平均错误率 ( $M_2 = 0.075$ ) 大于其他两种条件下 ( $M_1 = 0.049$ ,  $M_3 = 0.046$ ) 的平均错误率，具有中等效应；两者的交互作用不显著 ( $F(2,50) = 0.491$ ,  $p > .10$ )。

#### 3.2 不同字色条件随着反应任务的变化柱形图

根据 3.1 得到的平均数据，绘制柱形图如 3-2 所示，该图可以直观地发现识字任务的平均反应时高于辨色任务的平均反应时，但是差异不是很大；不同字色条件下的反应时以字色矛盾条件下的反应时最慢。

表 3-1 不同实验条件下的平均反应时 (ms) 和错误率

实验条件	辨色任务			识字任务		
	字色一致	字色矛盾	字色无关	字色一致	字色矛盾	字色无关
反应时	441.15±91.38	455.55±118.18	447.97±98.25	460.13±86.28	474.18±94.35	458.99±91.15
错误率	0.04±0.19	0.07±0.25	0.04±0.19	0.06±0.23	0.08±0.27	0.05±0.21

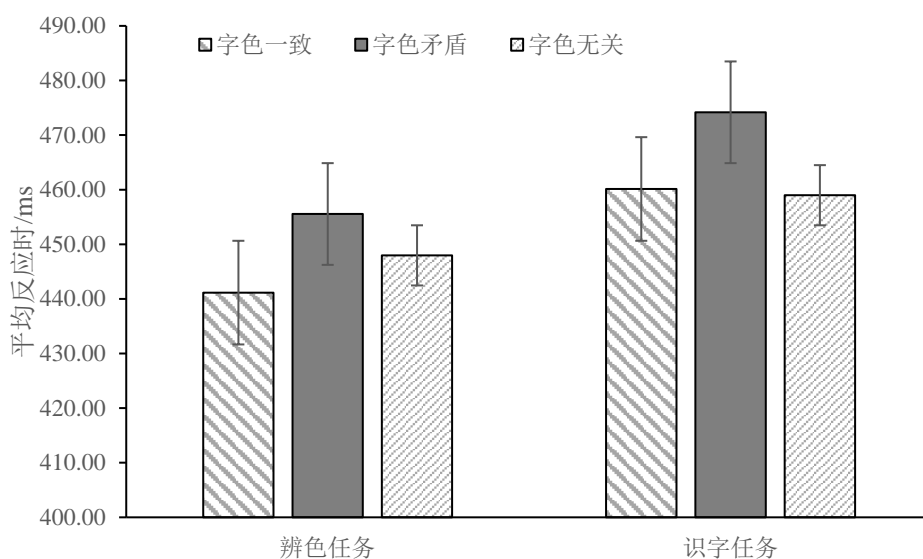


表 3-2 平均反应时在不同实验条件下的柱形图

## 4. 讨论

### 4.1 干扰现象分析

根据平均反应时数据和平均错误率数据可知,不同字色条件的主效应显著,字色矛盾下的平均反应时和错误率显著高于其他两种情况,说明字色矛盾情况严重干扰被试反应,即识字任务中颜色干扰识字,辨色任务中字义干扰辨色;并且实验任务的主效应显著,识字任务反应时较慢且错误率较大,说明识字任务下的字色不一致引起的干扰更大(但是效应并不是不强)。

将性别因素加入反应时的重复测量模型中,得到字色条件的主效应依旧是显著的( $F(2,24)=15.493$ ,  $p<.01$ ,  $\eta_p^2=0.564$ ),而发现性别的主效应不显著( $F(1,25)=0.124$ ,  $p>.10$ ),即不同性别被试的平均反应时不存在差异;性别因素与字色条件的交互作用不显著( $F(2,24)=0.671$ ,  $p>.10$ ),也就是说干扰现象不存在性别差异。将性别因素加入错误率的重复测量模型,得到字色条件的主效应依旧是显著的( $F(2,24)=9.312$ ,  $p<.01$ ,  $\eta_p^2=0.437$ ),发现性别的主效应不显著( $F(1,25)=0.136$ ,  $p>.10$ ),即不同性别被试的平均反应错误率不存在差异;性别因素与字色条件的交互作用不显著( $F(2,24)=1.743$ ,  $p>.10$ ),也就是说干扰现象不存在性别差异。综上可知,实验中存在 Stroop 效应(干扰现象),并且该干扰现象不存在性别差异。

### 4.2 与原 Stroop 结果差异分析

Stroop 在 1953 进行了三个实验对 Stroop 进行

研究,Exp1 中相当于是识字任务,结果发现被试对于字色不一致的单词的反应时比无关单词(black 颜色)要慢,反应时不存在无性别差异,但是存在年龄差异,分年级后发现大一大二的有干扰效应而大三大四的不存在,反应时差异上不存在性别差异和年级差异。Exp2 中相当于是个辨色任务,结果发现命名单词颜色和命名简单色块颜色之间存在差异,命名单词所需要的反应时大于简单色块,并且女生在颜色命名上相对于男生有优势,随着年级增大加工反应时更短,在反应时差异上存在一定的性别差异,不存在年级差异。Exp3 探究了练习效应,发现命名颜色不一致颜色时所需反应时随着练习下降迅速,存在这样的练习效应,在命名字色不一致的词语时速度反而降低。

本实验和原实验发现了字色不一致引起的干扰效应,并且该效应不存在显著的性别差异;并且练习效应对于两个实验任务的影响是不一致的。但是也存在一些不同之处,最大的不同之处在于原实验的识字任务反应时显著低于辨色反应时,识字任务的干扰效应(100 个字母的总反应时差异在 2s 左右)也显著小于辨色任务(100 个字母的总反应时差异在 40s 左右),而本实验却识字任务的平均反应时却大于辨色任务的平均反应时并且主效应显著,并且识字任务的 Stroop 效应和辨色任务中的 Stroop 效应不存在显著差异。

笔者认为可能存在以下原因:

(1) 任务本身不同。这一点体现在三个小点上,原实验的任务是单词,而本实验是汉字,两者对于被

试的干扰源是不一致的（刘海程，2007），使用不同语音的被试干扰作用不一样；原实验的任务计时是采用口头系列报告法，计算一百个单词的总反应时，而本实验采用的是计算机实验，考察的是每个汉字所花费的时间，两者在时间精度上也会有一定的差异；原实验以字色矛盾和字色无关的差异作为考察指标，但笔者在进行考察的时候并没有具体计算该效应指标，而是以字色条件的主效应作为衡量指标，并且；原实验只涉及两种实验条件，字色矛盾和字色无关，本实验还涉及到字色一致的条件，该条件的出现在一定程度或许会影响其他反应，原实验一定不存在字色一致，这对被试也存在一定引导作用。

（2）原实验选取了很多颜色，在颜色种类层面控制较好，而本实验只涉及两种颜色，相对单一，无法排除特定颜色加工速度差异等带来的潜在影响。

（3）原实验以字色矛盾和字色无关的差异作为考察指标，但笔者在进行考察的时候并没有具体计算该效应指标，而是以字色条件的主效应作为衡量指标，笔者的考察方式衡量的是三种实验条件之间的差异，而原实验只是两种实验条件之间的差异，或许会导致一定的差异。

以及被试选取也可能造成一定影响，因为我们的被试对于实验目的都是知道的。更多的差异可能还需要进一步的文献参考和实验研究。

#### 4.3 Stroop 效应的顺序效应和练习效应分析

字色条件的主效应即表示 Stroop 效应，Stroop 效应的顺序效应即研究实验顺序和字色条件的交互作用。对字色条件和实验顺序进行两因素的重复测量方差分析，可知实验顺序和交互作用的球形检验不符合（ $p>.05$ ），字色条件的球形检验符合，分别采用合适的检验结果可知，依然有字色条件的主效应显著，实验顺序的主效应不显著（ $F(1,25)=0.060$ ,

$p>.10$ ），即实验顺序不会显著影响判断反应时，实验顺序与字色条件的交互作用显著（ $F(2,24)=0.791$ ,  $p>.10$ ），即 Stroop 效应的顺序效应不显著。

分别在两个实验任务中以每 40 试次分为一个部分，分为三个区组，以区组和字色条件为自变量进行两因素重复测量考察两者交互作用的主效应，研究 Stroop 效应的练习效应即研究字色条件和区组之间的交互作用的主效应。对于辨色任务，球形检验符合（ $p>.05$ ），采用主体内检验，有字色条件主效应显著，即 Stroop 效应显著，字色条件与区组的交互作用不显著（ $F(4,204)=2.105$ ,  $p>.05$ ），说明练习效应不显著；对于识字任务，有字色条件的主效应显著，即 Stroop 效应显著，字色条件与区组的交互作用不显著（ $F(4,204)=9.312$ ,  $p<.05$ ,  $\eta_p^2=0.237$ ），说明练习效应不显著。综上可知，在辨色任务中，存在这样的练习效应，而在识字任务中不存在这样的练习效应。

#### 4.4 Stroop 效应影响因素

很长一段时间以来有许多关于 Stroop 效应的研究，但研究结果不尽相同，存在许多的影响因素会干扰 Stroop 效应，影响 Stroop 任务中对词语的加工。以下结合相关文献和本研究当中的结果进行分析

##### 4.4.1 颜色自身种类特性

有研究（Laeng，2005）表明当对被试进行两种不同类型颜色搭配的实验时，结果存在明显差异。在 Stroop 实验中，一个字义颜色，一个字本身颜色，他们根据色觉理论中的对立过程理论引入对立颜色（参见图 4-4-1-1），如果字义颜色和字本身颜色为对立的颜色（诸如黑白、红绿、黄蓝等鲜明对比的，例如黑色的“白”字），引起的 Stroop 干扰要小于字义颜色和字本身颜色不对立（诸如红蓝，例如红色的“蓝”字）的情况。

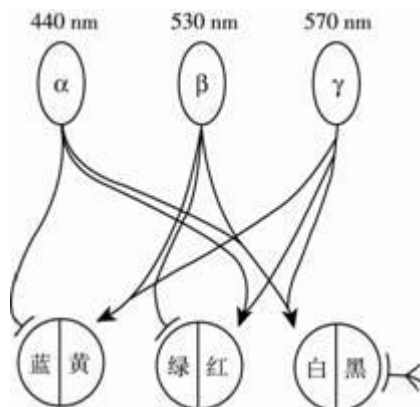


图 4-4-1-1 改编自心理学导论课本-色觉理论

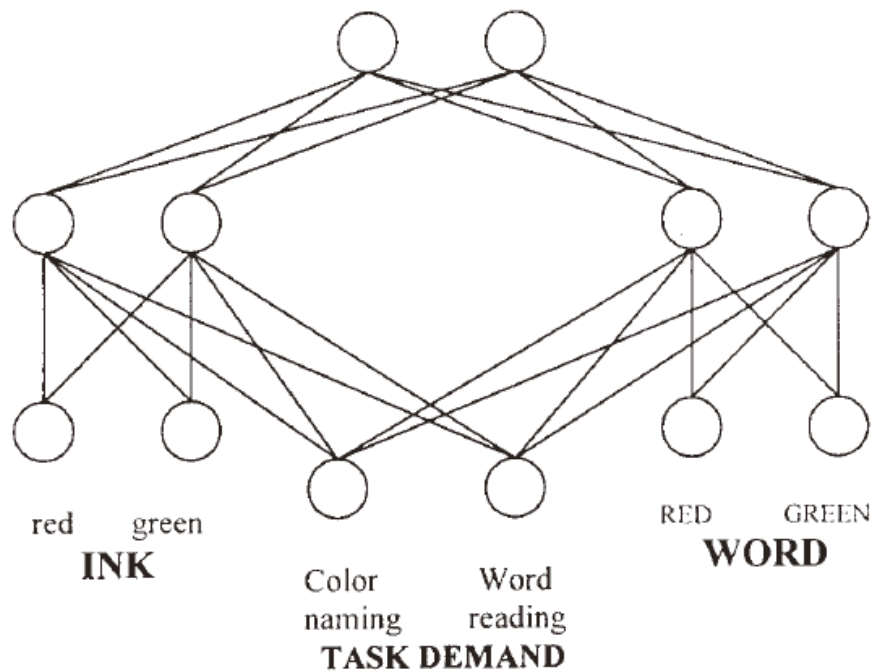


图 4-4-1-2 改编自 Cohen (1990)

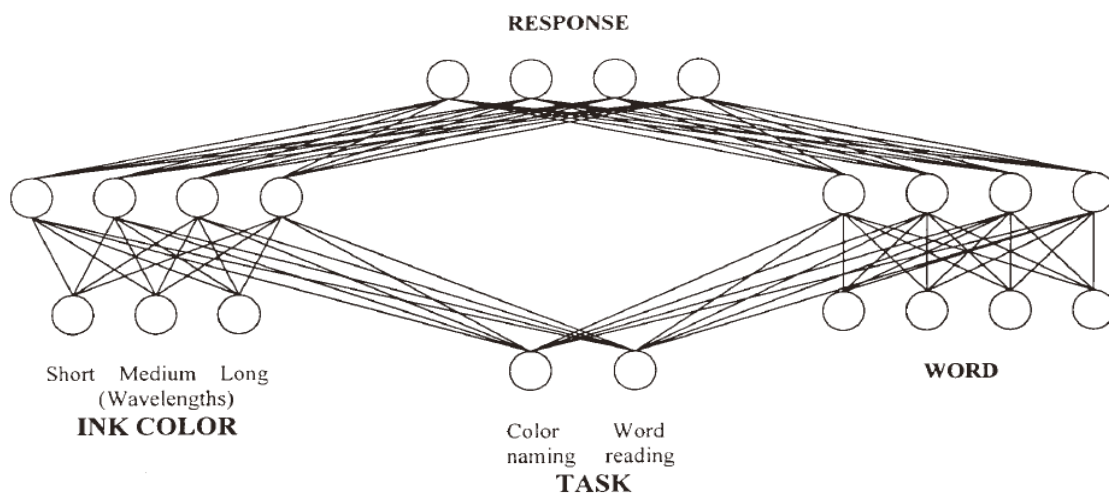


图 4-4-1-3 取自 Laeng et al (2005)

对此作者在 Cohen 的神经网络模型基础上改进并提出了自己的理论进行解释。Cohen 模型中的加工通路认为被试最后的判断受到字义和颜色的双通道输入影响（参见图 4-4-1-2），而作者则将这一点修改为波长因素，颜色本身的波长因素和颜色的含义会影响被试的加工和反应

#### 4.4.2 被试对颜色识别能力差异

有研究表明（Laeng, 2005）被试对颜色识别能力也会影响 Stroop 效应，研究引入标准的颜色视觉测试，发现在颜色视觉测试中表现错误多的被试在

Stroop 任务中的干扰效应也更大，两者之间呈现显著正相关。

#### 4.4.3 空间注意和任务要求的影响

Risko 等把 Stroop 范式和视觉搜索范式结合起来，研究发现词语的视觉加工要受到空间注意和任务要求影响，并不是大家认为的词加工是自动化的、被试不可避免要对词进行加工。Risko 给被试呈现 3、5、7 个词语，包括一个有颜色的靶子（其中实验 1 和 3 中靶子是词，实验 2 中靶子是一条有颜色的线）。不相关的颜色词也出现在呈现材料中，并且和靶子



的位置有两种情况：整体的、分离的。实验 1 和 3 中的任务是要求被试报告靶子的颜色，实验 2 的任务是要求被试确定靶子的颜色是否出现，结果在实验 1 当颜色词语和靶子是一个整体的时候产生了 Stroop 效应，但是在分离的情况下却没有产生 Stroop 效应。可见视觉词语的识别依赖于空间注意，空间注意的不确定性增加在一定程度上减少了外周词语被加工的可能性。

#### 4.4.3 语境因素

早在 80 年代，Smith 和 Thedor 等人就提出语境和加工深度之间的相关。随后 Simth 等通过实验证明语境能够调节语义的启动，并指出这是对自动化加工的挑战。Benser 等提出词语的视觉识别不是自动的，语境强烈影响并制约着词的加工。当处于注视点的中性词中所有字母都是有颜色的时候，一致条件下的反应时显著快于不一致条件的反应时；但是当中性词中仅仅有一个字母是有颜色的时候，一致和不一致条件下的反应时没有显著差异。Risko 和 Jennifer 等指出，对语义的加工存在严重的语境依赖性。据此我们提出语境或背景的强弱会影响 Stroop 效应加工的数量。背景复杂性越高干扰就会越大。如红蓝两个字之间呈现绿字，让被试命名中间绿字的颜色，这样颜色的干扰就比较大，Stroop 干扰的数量就越大。

Tulving 等认为在对词的加工中，遵循编码特异性原则，即编码和提取依赖于语境。Besner 和 Stolz 等曾报告他们的若干实验：在注视点呈现中性词语，用白色书写的颜色词呈现在中性词的上方或下方。其中中性词是有颜色的，要求被试报告中性词呈现的颜色，并且忽略呈现材料上其它词语。当注视点上中性词所有字母都是有颜色的时候，一致的条件显著快于不一致的条件；但是词中仅仅有一个字母是有颜色的时候，在一致和不一致条件之间没有显著不同。随后他们又进行了四个实验，进一步提出当词被看作非语言对象时候，在 Stroop 范式中仍存在严重的语境依赖性。据此，Besner 等提出加工的相对自动化假设，即便当词被看作非语言对象时候，语境对词的加工仍有强烈的影响。尽管以上两个研究都认为对词汇语义的加工，存在默认的加工模式，而非加工的完全自动化。但他们也承认该模式的发生机制有待探究。他们预言，随着认知神经科学的发展，特别像 fMRI、ERP 等技术的应用会给他们的观点提供支持。

通过以上陈述的研究者的发现我们能够分析，

本实验中并没有向被试提供一个语境，因此被试的语义编码并没有启动，自动化加工理论认为字义加工是自动化的，而颜色加工是控制加工，因此字义能够影响颜色辨识而反之则不能。但在本实验中并没有提供进行字义编码的语境，因此个体并未对字义进行编码，也就会出现本实验中的反转 Stroop 效应。

#### 4.4.4 年龄因素

研究发现 Stroop 干扰效应在小学低年级就已经出现，随着阅读技能的提高，大约在二三年级之后达到最高水平。随着阅读能力的继续发展，成年时期干扰会逐渐下降，直到接近 60 岁，此时干扰又会重新升高。彭冉龄（彭聃龄，2004）等采用事件相关电位技术，研究小学六年级 Stroop 效应的时间进程，与成人的 ERP 进行比较，发现儿童的 Stroop 效应产生机制可能与成人有一定的差异，认为儿童的 Stroop 效应的产生是一个复杂的过程，干扰效应是早期刺激颜色识别和后期反应选择两个阶段共同选择作用的结果。

### 4.5 Stroop 效应反转和情绪 Stroop 任务

#### 4.5.1 Stroop 效应反转

Stroop 效应的反转范式是由 Logan 提出的，后由 Merikle 等人对其改进。Merikle 等人利用 Stroop 色词干扰任务的一种变式进行了有意思知觉和无意识知觉的研究。这种实验范式最早由 Logan 等人[9]提出，后来被 Merikle 等人用于无意识知觉的研究，它与传统的 Stroop 任务的实验范式不同，只涉及红、绿两种颜色，色词“RED”或“GREEN”是启动刺激，用于启动对红、绿两种靶颜色的命名反应。在实验中首先给被试呈现一个灰色的色词，如 RED 或 GREEN，对色词的知觉或者是有意识的（呈现时间较长），或者是无意识的（呈现时间较短）。接着呈现一个色块（红色或绿色）让被试尽可能快地命名。色词和色块有时一致（如 RED 和红色块），有时不一致（如 RED 和绿色块）。当一致的色词—色块对发生的概率（如 25%）远小于不一致发生的概率（如 75%）时，被试的反应结果依赖于对色词的知觉是有意识的还是无意识的。在意识状态下，被试对不一致的色块的命名要快于对一致的色块的命名，出现了典型的 Stroop 效应的反转。而在无意识状态下，被试对一致的色块的命名要快于对不一致的色块的命名，出现了典型的 Stroop 效应。

综上，简单地讲，Stroop 效应的反转指的是被试对不一致色块的命名反而快于字色一致色块的命



名，即原本设置可能的冲突变成了促进效应。

#### 4.5.2 情绪 Stroop 任务

情绪 Stroop 效应最初是在研究有情绪障碍的个体时发现的，它是指，当命名用不同颜色写成的“情绪词”和“中性词”的颜色时，前者的颜色命名时间要长于后者。这种对自己当前关注或者说是与其情绪相关的词作颜色命名判断时，所出现的干扰效应就是情绪 Stroop 效应。关于情绪刺激对注意影响的研究，经常采用情绪 Stroop 范式。在情绪 Stroop 任务中，通过呈现给被试一些情绪刺激物，要求被试对这些刺激物中的非情绪信息做出快速反应，以考察被试对情绪刺激的注意偏向及对非情绪信息认知加工时的抗情绪干扰情况。以情绪词或情绪面孔颜色命名的反应时与非情绪词或中性面孔颜色命名的反应时之差，来代表情绪信息所产生的注意偏向：如果反应时之差为正数，表明词或面孔的情绪信息得到了加工，干扰了对颜色的命名。

情绪 Stroop 效应广泛运用于心理学各个领域，目前发现：不同症状的被试对与其病症有关的情绪词所产生的干扰最大；如果实验材料与被试当前所关注的事物有关的话，所产生的干扰就更大；不仅成人被试，患病的儿童也发现了情绪 Stroop 效应；刺激呈现时间不同会对该效应造成影响；情绪特质

和情绪状态也会对该效应造成影响。

#### 4.6 Stroop 效应范式的实验变式及其应用

除了前面的 Stroop 效应反转范式和情绪 Stroop 任务，经典 Stroop 范式还有以下几个反应变式

##### 4.6.1 昼夜 Stroop 任务范式 (Day-Night Stroop)

Cerstadt、Hong 和 Diamond 等修订了 Stroop 任务 (Gerstadt et al, 1994)，使之适用于 3.5 岁至 7 岁的学前儿童，形成昼夜 Stroop 任务 (参见流程图 4-6-1)。实验组要求被试在看见太阳的图形时说“夜晚”，看见月亮的图形时说“白天”；控制组要求被试看见一个抽象的图形时说“白天”，看见另一个抽象图形时说“夜晚”。结果实验组和控制组之间有显著差异。实验组被试的反应时有显著的年龄差异，尤其是在 3.5~4.5 岁之间。控制组的被试反应时则没有表现出显著的年龄差异。被试反应时随着试验次数增多而减少。最普遍的错误模式是儿童会交替说“白天”和“夜晚”，或看见太阳的轮廓坚持说“白天”。昼夜 Stroop 和经典 Stroop 相同性在于两者都是要求儿童抑制字意义和视觉冲突的矛盾，但是前者同时又是一个规则性任务。因为实验任务是要求被试见太阳说“夜晚”，而看见月亮或者星星说“白天”。当前，昼夜 Stroop 范式大多被应用于研究儿童的执行功能，作为测量执行功能的一个重要指标。

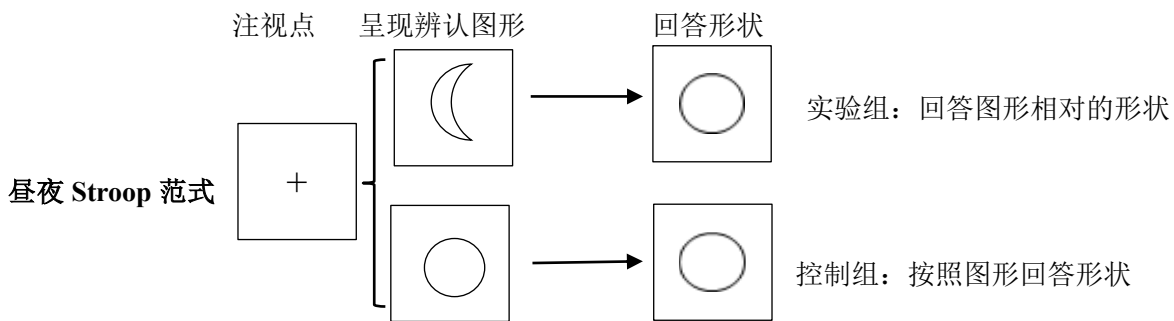


图 4-6-1 Day-Night Stroop 流程示意图

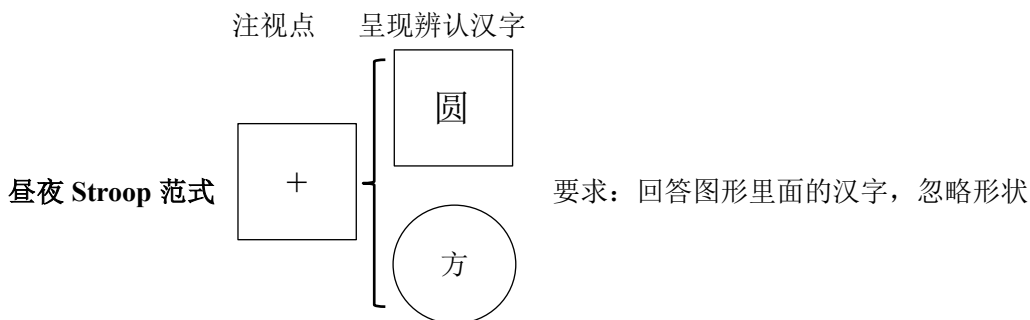


图 4-6-2-1 图-词干扰范式流程示意图

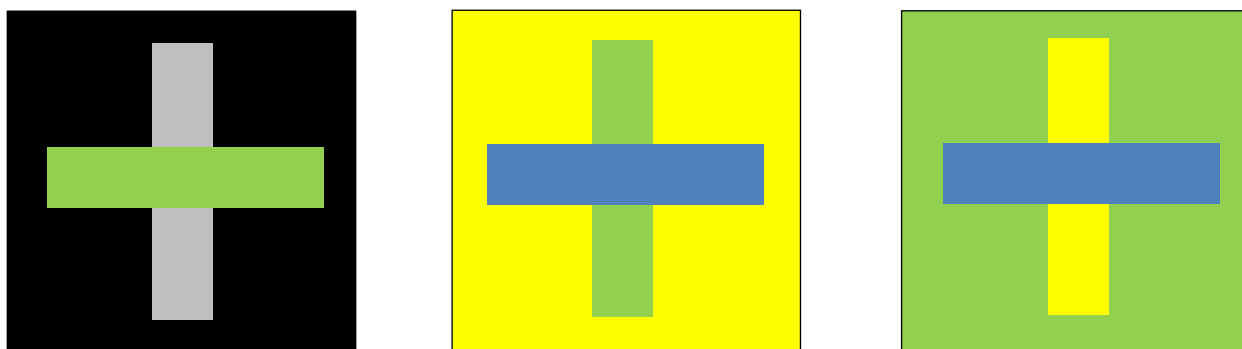


图 4-6-2-2 图词干扰范式流程改进

#### 4.6.2 图-词干扰范式

图-词干扰任务是经典 Stroop 范式的一个简单推广。在图-词干扰范式中，向被试呈现图与干扰词，二者同时呈现，或者先呈现图，然后呈现干扰词，要求被试命名干扰词。例如向被试呈现圆形中的汉字“方”、正方形中“圆”字等，任务是要求被试忽略图形，命名图形里面的汉字。对于图-词干扰范式，近年来研究者对其作进一步改进，Waszak 和 Martina 等人（Waszak, 2003; Martina, 2005）让被试命名两个长方形其中一个的颜色，颜色字所在的位置有三种情况：在靶子目标上、非靶子目标、背景上，结果发现在靶子目标上的颜色字比在非靶子目标上和背景上的颜色字产生较大的 Stroop 效应（参见图 4-6-2-2）。另外，图词-干扰范式也是研究双 Stroop 效应的重要途径和方法

#### 4.6.3 双语 Stroop 范式

双语 Stroop 范式指利用两种语言的色词，要求被试用母语和第二语言分别对两种语言的色词进行颜色命名，然后根据语言间（命名语言和色词使用两种语言）和语言内（命名语言和色词使用同一种语言）的不同干扰效果推论出双语者的心理词典表征结构。翻译任务也是研究双语 Stroop 范式的重要工具。Laheij 和 Debruyne 等采用翻译任务检验双语中 Stroop 效应。在实验中要求被试进行从第二语言到第一语言的翻译，干扰词则以第一语言呈现，并且要求干扰词与要生成词语在正字法、语义等方面存在相关。结果在干扰词呈现 140ms 的时候，与生成的词语语义上相关的干扰词产生干扰作用，翻译词语的反应时显著增加，而与生成的词在正字法上相关的干扰词促进了翻译任务的完成，反应时显著缩短。这正如同经典 Stroop 效应中的不一致条件下的干扰和一致条件下促进作用。

## 5. 结论

本实验通过中文汉字的词义干扰 Stroop 范式对 Stroop 效应进行探究，结果发现：（1）当颜色和字义两个维度不一致的时候，会对被试的反应产生干扰，并且两个实验任务之间的 Stroop 效应没有显著差异（2）Stroop 不存在性别差异，也不存在顺序效应，（3）辨色任务中的 Stroop 效应不存在明显的练习效应，识字任务中的 Stroop 效应存在一定的练习效应（4）Stroop 效应受到诸多因素，例如颜色字义搭配、颜色识别能力、语境因素、年龄因素等。

## 6. 思考

### 6.1 两种任务之间的 Stroop 效应分析

按照原本的实验研究表明，应当是字色对字义的辨别几乎没有什么影响，而字义对字色辨别有显著影响，但是在本实验中都发现了 Stroop 效应；即便如此，按照字义是自动化加工，而颜色是控制加工的理论解释，也应该是辨色任务中的 Stroop 干扰效应应当大于识字任务汇总的 Stroop 效应，而本实验当中出现了倒置（参见图 3-2）。

字色条件的主效应代表 Stroop 效应，字色条件与实验任务之间的交互代表了 Stroop 效应在两种任务之间的差异，实际上方差分析检验时不显著的（ $F(2,50) = 2.237, p > .10$ ），这一点在前面已经提到，也就是说两种任务之间的差异并不显著，由此可见自动化理论并不能很好解释实验结果，基于神经网络的平行分布式加工模型对本实验的结果可以更好解释。

观察图 3-2，发现识字任务的平均反应时大于辨色任务平均反应时，这一点方差检验也是显著的，即实验任务主效应显著。刚刚已经分析得出 Stroop

效应在两个任务之间的差异不显著，所以该平均反应时的差异来自固有差异，即加工反应时，可见识字速度并非快于颜色加工速度，反而可能是识字速度慢于颜色加工速度，这点差异来源尚且不清楚。

## 6.2 汉字、拼音、字母的 Stroop 效应

在前面刚刚提到识字加工速度或许并非快于颜色加工速度，当然也可能是被试选取差异。如果从加工速度角度出发，我们可以联想到汉字、拼音、字母之间或许会存在差异，对于中国被试来讲，汉字加工速度可以很快，但是拼音加工速度较慢，英文

加工速度更慢（英语专业学生或许对英语加工快于拼音），因此这一点差异值得研究。

查阅文献发现已有人（刘海程，2007）对该问题进行研究，研究结果发现汉字干扰大于拼音大于字母，研究认为汉字的干扰来自语音和语义两个方面，拼音干扰则主要来自语音，英文字母相对于汉字和拼音，主要来源于语义。汉字的干扰源多于拼音和字母，所以干扰效应最大，作者利用文字加工过程的平行分布加工模型对该结果进行了很好的解释。

### 参考文献

- Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: performance of children, ja:math-7 years old on a stroop- like day-night test. *Cognition*, 53(2), 129-53.
- Wühr, P., & Waszak, F. (2003). Object-based attentional selection can modulate the stroop effect. *Memory & Cognition*, 31(6), 983-94.
- Wühr, P., & Welte, M. (2005). Depth cues do not underlie attentional modulations of the stroop effect. *Memory & Cognition*, 33(4), 676-680.
- Laeng, B., Låg, T., & Brennen, T. (2005). Reduced stroop interference for opponent colors may be due to input factors: evidence from individual differences and a neural network simulation. *Journal of Experimental Psychology Human Perception & Performance*, 31(3), 438-52.
- Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology General*, 121(1), 15-23.
- Macleod, C. M. (1991). Half a century of research on the stroop effect: an integrative review. *Psychological Bulletin*, 109(2), 163-203.
- Besner, D., & Stolz, J. (1999). Context dependency in stroop's paradigm: when are words treated as nonlinguistic objects?. *Canadian Journal of Experimental Psychology*, 53(4), 374-80.
- 彭聃龄, 郭桃梅, 魏景汉, & 肖丽辉. (2004). 儿童 stroop 效应加工阶段特点的事件相关电位研究. *科学技术与工程*, 4(2), 84-88.
- 刘海程, & 翁旭初. (2007). 汉字、拼音、英文 stroop 干扰效应的比较研究. *心理科学*, 30(2), 365-368.
- 鞠鑫, & 刘鸣. (2003). 双语 stroop 效应研究范式的探讨. *社会心理学*(4), 43-46.
- 耿海燕, & 朱滢. (2001). Stroop 效应及其反转:无意识和意识知觉. *心理科学*, 24(5), 553-556.
- 陈俊, 刘海燕, & 张积家. (2007). Stroop 效应研究的新进展——理论、范式及影响因素. *心理科学*, 30(2), 415-418.

## 附录

表 1 所有被试不同识字任务和辨色任务下的平均反应时

姓名	辨色任务			识字任务		
	字色一致	字色矛盾	字色无关	字色一致	字色矛盾	字色无关
蔡贝贝	387.48	405.13	387.15	446.55	433.48	445.20
岑梦媛	491.35	470.18	489.83	463.75	470.10	450.08
陈滋正	443.98	419.65	448.53	499.28	526.85	523.93
樊丹红	464.25	534.73	475.95	478.10	475.83	471.55
方略	463.60	463.85	456.90	472.45	552.95	536.33
方其为	388.85	376.28	393.93	413.78	423.83	401.78
符凯优	430.38	440.88	437.93	508.93	505.25	502.53
郭佳	502.85	477.28	514.40	478.40	490.70	458.33
胡沁艺	437.80	426.80	455.48	449.08	458.08	454.93
黄杨初	507.98	564.13	510.28	417.13	457.53	465.90
蒋瑞峰	393.70	391.55	393.73	448.35	436.10	423.98
金多	437.93	429.85	421.35	477.60	482.40	469.03
黎松	487.15	498.75	505.33	534.25	513.85	504.88
李伟	432.68	432.75	439.05	475.48	524.53	504.20
李炜	449.85	438.38	457.18	518.35	531.90	473.00
李蔚	431.15	574.73	441.68	509.03	481.60	492.58
李文敏	452.80	474.25	462.93	502.85	550.53	499.95
李彦哲	421.55	420.38	417.55	405.53	406.13	414.13
梁展鸿	406.10	410.33	404.10	422.68	453.53	418.75
梁子悦	459.05	543.60	480.53	469.90	490.88	472.08
廖一帆	533.25	551.95	530.38	519.33	525.23	508.43
刘嘉耀	489.90	507.10	525.25	514.43	572.70	528.75
刘锡隆	398.70	424.80	439.83	446.48	461.78	450.60
刘子薇	357.10	335.28	352.65	375.60	412.83	377.63
娄佳飞	428.03	446.63	401.80	473.08	475.70	465.25
卢家乐	380.88	379.70	394.43	487.75	470.30	482.93
陆辰馨	416.80	415.70	434.48	423.70	427.58	454.68
马一诺	416.65	444.03	454.95	381.30	378.85	367.60
任帅	403.10	406.65	408.53	424.55	438.48	432.85
盛超悦	454.70	446.18	445.20	449.23	487.50	456.55
施姚希	466.08	468.35	455.78	468.05	456.40	468.48
宋津南	452.45	460.18	491.45	473.98	502.30	514.15
万子薇	474.38	461.30	463.75	531.03	547.18	519.30
汪明慧	519.88	556.90	517.50	458.78	483.00	465.43
王力	374.35	390.83	375.90	417.58	426.13	431.38
王曼华	435.93	405.40	428.45	466.15	492.95	450.55
王晓君	424.75	449.18	432.23	454.75	463.20	443.23
王泽振	421.75	425.78	415.70	438.83	458.25	445.05
温馨	520.00	610.95	550.00	478.40	527.03	506.43
吴思远	396.78	451.18	409.25	459.15	507.08	450.05

谢未佳	389.73	396.63	398.50	435.05	451.73	429.53
叶迪	434.85	446.43	430.90	443.50	495.10	472.20
俞一凡	395.63	428.48	433.10	420.90	398.88	391.13
詹少炜	395.40	349.43	399.60	453.48	455.43	448.68
张珂萌	411.90	417.20	437.75	431.93	405.58	384.68
张莅宸	407.68	403.60	401.23	398.88	406.75	397.85
章嘉斐	503.73	562.33	508.95	522.90	533.58	514.30
赵剑辉	403.75	430.40	407.55	439.40	459.70	450.08
郑程如	451.88	478.78	464.10	460.23	476.78	446.20
郑林枫	574.25	625.00	588.55	427.35	445.40	412.35
朱静茵	450.58	434.50	443.48	466.73	484.55	441.55
朱月平	464.70	484.33	459.70	493.00	463.33	476.68
总计	441.15	455.55	447.97	460.13	474.18	458.99

表 1 所有被试不同识字任务和辨色任务下的平均错误率

姓名	辨色任务			识字任务		
	字色一致	字色矛盾	字色无关	字色一致	字色矛盾	字色无关
蔡贝贝	0.03	0.08	0.03	0.00	0.00	0.00
岑梦媛	0.00	0.05	0.03	0.05	0.08	0.05
陈滋正	0.13	0.03	0.08	0.25	0.13	0.15
樊丹红	0.00	0.00	0.00	0.03	0.00	0.00
方略	0.00	0.08	0.00	0.08	0.18	0.08
方其为	0.03	0.00	0.00	0.10	0.13	0.03
符凯优	0.00	0.05	0.03	0.10	0.05	0.05
郭佳	0.03	0.10	0.00	0.03	0.00	0.03
胡沁艺	0.00	0.18	0.00	0.03	0.05	0.05
黄杨初	0.00	0.05	0.03	0.03	0.03	0.03
蒋瑞峰	0.03	0.05	0.00	0.00	0.15	0.10
金多	0.03	0.05	0.03	0.00	0.10	0.08
黎松	0.10	0.15	0.10	0.05	0.15	0.08
李伟	0.00	0.03	0.00	0.00	0.10	0.03
李炜	0.10	0.00	0.08	0.10	0.08	0.18
李蔚	0.00	0.13	0.05	0.05	0.10	0.05
李文敏	0.05	0.05	0.13	0.00	0.10	0.08
李彦哲	0.08	0.05	0.03	0.08	0.08	0.18
梁展鸿	0.10	0.13	0.13	0.10	0.20	0.10
梁子悦	0.03	0.15	0.03	0.00	0.05	0.03
廖一帆	0.00	0.05	0.00	0.03	0.03	0.00
刘嘉耀	0.05	0.08	0.08	0.08	0.08	0.00
刘锡隆	0.05	0.05	0.05	0.05	0.05	0.08
刘子薇	0.05	0.05	0.05	0.13	0.05	0.08

娄佳飞	0.03	0.08	0.05	0.05	0.03	0.05
卢家乐	0.10	0.03	0.00	0.03	0.00	0.05
陆辰馨	0.00	0.05	0.03	0.03	0.08	0.00
马一诺	0.00	0.15	0.08	0.03	0.10	0.08
任帅	0.00	0.10	0.00	0.00	0.10	0.03
盛超悦	0.03	0.18	0.05	0.08	0.08	0.03
施姚希	0.13	0.15	0.08	0.15	0.10	0.10
宋津南	0.05	0.18	0.05	0.13	0.10	0.00
万子薇	0.08	0.00	0.05	0.03	0.10	0.08
汪明慧	0.00	0.05	0.10	0.03	0.03	0.03
王力	0.03	0.00	0.00	0.00	0.13	0.03
王曼华	0.08	0.00	0.00	0.05	0.03	0.10
王晓君	0.03	0.00	0.03	0.05	0.08	0.00
王泽振	0.00	0.00	0.00	0.08	0.08	0.03
温馨	0.05	0.15	0.05	0.03	0.03	0.00
吴思远	0.03	0.03	0.03	0.08	0.10	0.00
谢未佳	0.10	0.03	0.03	0.10	0.20	0.08
叶迪	0.03	0.00	0.00	0.08	0.20	0.05
俞一凡	0.13	0.08	0.10	0.08	0.03	0.10
詹少炜	0.08	0.10	0.10	0.10	0.25	0.08
张珂萌	0.00	0.00	0.03	0.03	0.00	0.00
张莅宸	0.05	0.03	0.00	0.10	0.05	0.03
章嘉斐	0.00	0.08	0.03	0.10	0.05	0.05
赵剑辉	0.03	0.03	0.03	0.00	0.03	0.03
郑程如	0.00	0.05	0.05	0.03	0.03	0.00
郑林枫	0.03	0.40	0.05	0.08	0.03	0.03
朱静茵	0.03	0.08	0.08	0.03	0.05	0.03
朱月平	0.05	0.05	0.03	0.10	0.03	0.03
总计	0.04	0.07	0.04	0.06	0.08	0.05