

# 基于序列感知的推荐算法研究

付海林

计算机学院  
华南师范大学

2019 年 5 月 28 日



# 目录

- ① 研究的背景与意义
- ② 文献综述
- ③ 研究方法 with 框架



# 研究背景与目的

- 现有推荐策略都没有考虑用户兴趣变化的动态性
- 工业界主流的点击率预估类推荐算法都需要构造大量的用户隐私相关的特征
- 在某些特征的场景下 (比如用户未登陆或处于匿名状态, 现有推荐算法将无法发挥作用)
- 现有的推荐策略基本都忽略了用户行为之间的时序关系



# 研究背景与目的

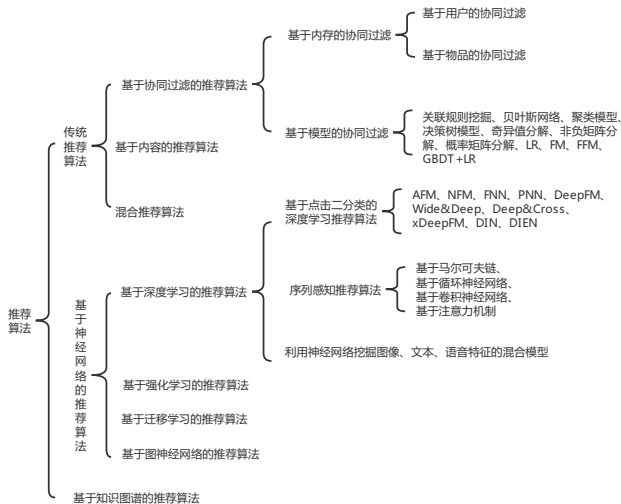
- 现有推荐策略都没有考虑用户兴趣变化的动态性
- 工业界主流的点击率预估类推荐算法都需要构造大量的用户隐私相关的特征
- 在某些特征的场景下 (比如用户未登陆或处于匿名状态, 现有推荐算法将无法发挥作用)
- 现有的推荐策略基本都忽略了用户行为之间的时序关系

## 研究目的

利用神经网络序列建模技术挖掘用户行为之间的时序关系



# 推荐算法分类思维导图



# 提高记忆效能的意义

- 有益于数学学习的减负增效
- 调动多重记忆系统有利于激活各种类型思维
- 多样化的记忆策略兼顾学生数学学习的个性化，助其扬长补短
- 形成正确的记忆观念和学习态度，去伪存真让思维焕发活力



# 提高记忆效能的意义

- 有益于数学学习的减负增效
- 调动多重记忆系统有利于激活各种类型思维
- 多样化的记忆策略兼顾学生数学学习的个性化，助其扬长补短
- 形成正确的记忆观念和学习态度，去伪存真让思维焕发活力



图: 搜索 App

基于序列感知的推荐算法研究

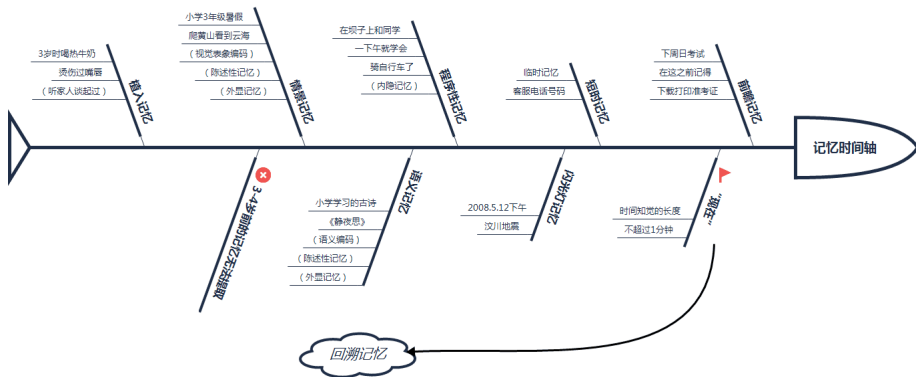


图: 时间轴上的记忆类型





# 记忆模型

记忆虽然就是通过感觉获得信息，进行加工 (将信息转换成可用的形式)，组织存储的信息，并从存储中提取信息的过程。但它绝不是简单的对情景“拍照”，对语义的“笔录”，而是一个建构重塑的过程。



# 记忆的模式

记忆虽然就是通过感觉获得信息，进行加工 (将信息转换成可用的形式)，组织存储的信息，并从存储中提取信息的过程。但它绝不是简单的对情景“拍照”，对语义的“笔录”，而是一个建构重塑的过程。

## 能够触及记忆本质的两个记忆模型

### 记忆多重存储模型

(得到脑科学研究临床病例实证支持，确有短时记忆转入长时记忆障碍患者，模型表明大脑中确实存在不同类型的记忆，揭示了信息被编码、保持、提取的过程)

### 记忆加工水平模型

(可以更好的解释对信息进行不同水平的加工会产生不同记忆效果的现象，还能很好的解释自我参照效应 (self-reference effect) 可以增强记忆效果)



# 记忆生理机制

成人几乎所有部位的神经元是不能再生更新的，到是神经元的突触连接是在增加的，故而智力在不断发展。海马回中的齿状回，是成人能持续产生新神经元的少数几个区域之一。这个区域已经被证明是形成记忆的重要区域。



# 记忆生理机制

成人几乎所有部位的神经元是不能再生更新的，到是神经元的突触连接是在增加的，故而智力在不断发展。海马回中的齿状回，是成人能持续产生新神经元的少数几个区域之一。这个区域已经被证明是形成记忆的重要区域。

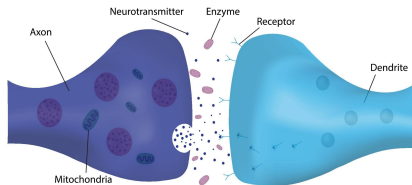
## 记忆如何存储与存储在哪里？

短时存储表现为神经递质的释放增加，长时存储表现为新突触的生成(突触是一个神经元的轴突与另一个神经元的树突之间的狭窄间隙。)



# Memory

## Poor Memory



## Good Memory

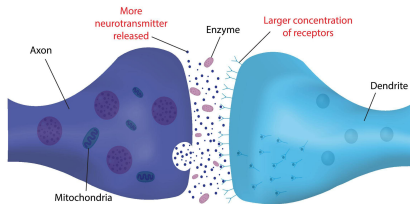


图: 神经递质释放增加促成短时存储



# 工作记忆

工作记忆掌控着意识流与思维走向，它不只是简单地你记住了什么，而是你可以以怎样的效率干什么。

工作记忆还包含诸如连接、转换等信息处理过程，仅凭短时记忆存储是无法实现的。



# 工作记忆

工作记忆掌控着意识流与思维走向，它不只是简单地你记住了什么，而是你可以以怎样的效率干什么。

工作记忆还包含诸如连接、转换等信息处理过程，仅凭短时记忆存储是无法实现的。

**语音回路** 主要用于记住词的顺序，保持信息

**视觉空间模板** 主要用于加工视觉和空间信息 (比如心理旋转)

**中央执行系统** 除了控制加工以外也有一定信息储存能力  
(调控中如果某一子系统需要, 在调用其他子系统资源进行支援时用作为中转储存)



# 工作记忆模型

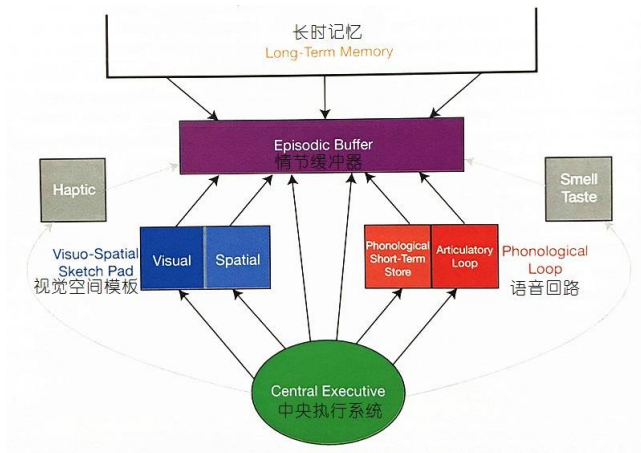


图: 改进后的 Baddeley 工作记忆模型



# 长时记忆

- 实验证实短时记忆和长时记忆遗忘的主要原因均是干扰 (曾一度认为短时记忆遗忘的主要原因是消退)
- “干扰抑制回忆, 提取诱发遗忘” 这个研究结果对数学教学很有意义
- 长时记忆既包含陈述性记忆, 还包含程序性记忆
- 长时记忆的组织方式与因特网类似, 储存的知识具有语义网络的结构, 语义相关的都会产生超链接
- 通过训练和长期实践应用特定领域的记忆技能可以形成长时工作记忆, 记忆中的专家效应



# 记忆研究新领域

**元记忆** 监控个体自身记忆对象和进程的记忆

**自传体记忆** 伴有情绪感情色彩，构建自身基模之上，附着个人个性信息的记忆

**内隐记忆** 无需意识参与的记忆

**前瞻记忆** 指向未来带有计划性的记忆

**错误记忆** 被暗示，产生偏差，被动植入等导致的有悖事实的记忆



# 研究方法

## 文献研究

对于脑科学和心理学，自身专业知识和实验条件均存在短板，主要通过阅读学习相关学术文献，试图将架构更底层，具备一般性的前沿研究成果运用到中学数学课堂教学中



# 研究方法

## 文献研究

对于脑科学和心理学，自身专业知识和实验条件均存在短板，主要通过阅读学习相关学术文献，试图将架构更底层，具备一般性的前沿研究成果运用到中学数学课堂教学中

## 案例研究

对中学数学课堂教学中积累的素材、典型案例、经验做法回顾反思，收集有助于记忆的教学素材，提炼有效的经验做法



# 研究方法

## 行动研究

对已有一定教学积累，感觉有效的记忆策略，且在以往教学中得到部分验证的想法、课例，展开行动研究。重新思考教学计划、展开行动、多方面观察教学实效，在此基础上进行反思归纳。更新想法，再修订，再展开新一轮行动研究



# 研究方法

## 行动研究

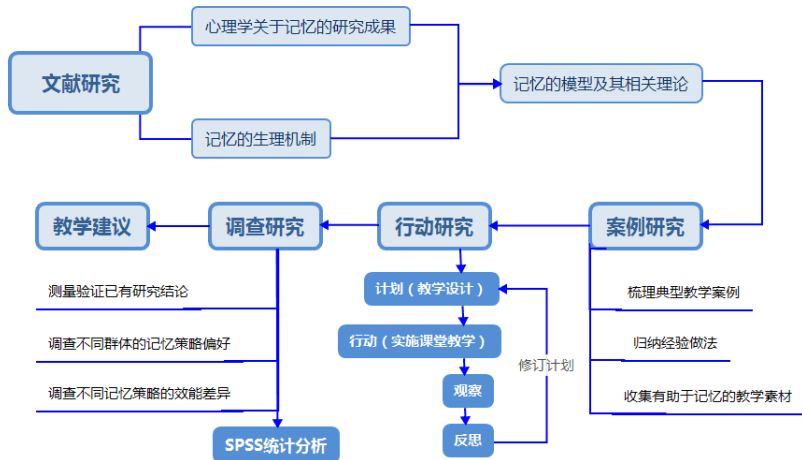
对已有一定教学积累，感觉有效的记忆策略，且在以往教学中得到部分验证的想法、课例，展开行动研究。重新思考教学计划、展开行动、多方面观察教学实效，在此基础上进行反思归纳。更新想法，再修订，再展开新一轮行动研究

## 调查研究

尝试验证文献研究中结论有出入、观点有分歧的问题；了解元记忆的监测与控制水平；验证之前一些仅凭经验归纳提炼的数学记忆策略



# 研究框架



## 第一部分

### 第一节

请插入图片

图:





# 提高工作记忆活动效率的策略

## 提高信息块 (chunks) 的质量

$$\begin{aligned}x^3 + y^3 + z^3 - 3xyz &= \\(x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx) \\ \sum x^3 - 3xyz &= (\sum x)(\sum x^2 - \sum xy)\end{aligned}$$



# 提高工作记忆活动效率的策略

## 提高信息块 (chunks) 的质量

$$\begin{aligned}x^3 + y^3 + z^3 - 3xyz &= \\(x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx) \\ \sum x^3 - 3xyz &= (\sum x)(\sum x^2 - \sum xy)\end{aligned}$$

## 优化算法减轻信息块调用负荷

$$74 \times 76 \longrightarrow \overline{7 \times 8} \quad \overline{4 \times 6} \longrightarrow \overline{5624}$$



# 提高工作记忆活动效率的策略

## 提高信息块 (chunks) 的质量

$$\begin{aligned}x^3 + y^3 + z^3 - 3xyz &= \\(x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx) \\ \sum x^3 - 3xyz &= (\sum x)(\sum x^2 - \sum xy)\end{aligned}$$

## 优化算法减轻信息块调用负荷

$$74 \times 76 \longrightarrow \overline{7 \times 8} \quad \overline{4 \times 6} \longrightarrow \overline{5624}$$

## 发展长时工作记忆，形成专家记忆优势效应

Basel 问题  $1 + \frac{1}{4} + \frac{1}{9} \cdots + \frac{1}{k^2} + \cdots = \frac{\pi^2}{6}$   
欧拉计算时毫不费力，就像人呼吸或者鹰在风中  
保持平衡一样——物理学家阿拉果 (法)



# 提高工作记忆活动效率的策略

## 提高信息块 (chunks) 的质量

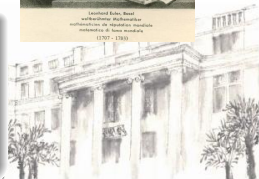
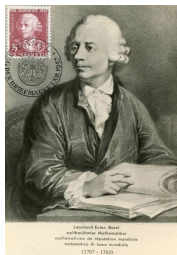
$$\begin{aligned}x^3 + y^3 + z^3 - 3xyz &= \\(x + y + z)(x^2 + y^2 + z^2 - xy - yz - zx) \\ \sum x^3 - 3xyz &= (\sum x)(\sum x^2 - \sum xy)\end{aligned}$$

## 优化算法减轻信息块调用负荷

$$74 \times 76 \longrightarrow \overline{7 \times 8} \quad \overline{4 \times 6} \longrightarrow \overline{5624}$$

## 发展长时工作记忆，形成专家记忆优势效应

Basel 问题  $1 + \frac{1}{4} + \frac{1}{9} \cdots + \frac{1}{k^2} + \cdots = \frac{\pi^2}{6}$   
欧拉计算时毫不费力，就像人呼吸或者鹰在风中保持平衡一样——物理学家阿拉果 (法)



# 提升长时记忆效果的策略

## 及时回顾与形式丰富的回忆

$$\begin{aligned}
 \sin \alpha &= \frac{y}{r} = \frac{y}{\sqrt{x^2+y^2}} = \frac{1}{\csc \alpha} = \pm \frac{1}{\sqrt{1+\cot^2 \alpha}} \\
 &= \tan \alpha \cos \alpha = \frac{\cos \alpha}{\cot \alpha} = \pm \sqrt{1 - \cos^2 \alpha} \\
 &= \cos\left(\frac{\pi}{2} - \alpha\right) = -\cos\left(\frac{\pi}{2} + \alpha\right) = \sin(\pi - \alpha) = -\sin(\pi + \alpha) \\
 &= -\cos\left(\frac{3\pi}{2} - \alpha\right) = \cos\left(\frac{3\pi}{2} + \alpha\right) = -\sin(-\alpha) = \sin(2k\pi + \alpha) \\
 &= \sin(\alpha - \beta) \cos \beta + \cos(\alpha - \beta) \sin \beta = \sin(\alpha + \beta) \cos \beta - \cos(\alpha + \beta) \sin \beta \\
 &= \sin\left(\frac{\alpha+\beta}{2}\right) \cos\left(\frac{\alpha-\beta}{2}\right) + \cos\left(\frac{\alpha+\beta}{2}\right) \sin\left(\frac{\alpha-\beta}{2}\right) \\
 &= \sin\left(\frac{\beta+\alpha}{2}\right) \cos\left(\frac{\beta-\alpha}{2}\right) - \cos\left(\frac{\beta+\alpha}{2}\right) \sin\left(\frac{\beta-\alpha}{2}\right) \\
 &= 2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2} = (\sin \frac{\alpha}{2} + \cos \frac{\alpha}{2})^2 - 1 = 1 - (\sin \frac{\alpha}{2} + \cos \frac{\alpha}{2})^2 = \frac{2 \tan \frac{\alpha}{2}}{1 + \tan^2 \frac{\alpha}{2}} \\
 &= \frac{2}{\tan \frac{\alpha}{2} + \cot \frac{\alpha}{2}} = \pm \sqrt{\frac{1 - \cos 2\alpha}{2}} = 3 \sin \frac{\alpha}{3} - 4 \sin^3 \frac{\alpha}{3} \\
 &= \sqrt{2} \sin\left(\alpha + \frac{\pi}{4}\right) - \cos \alpha = \dots
 \end{aligned}$$

# 提升长时记忆效果的策略

## 优化编码方式深度加工记忆信息

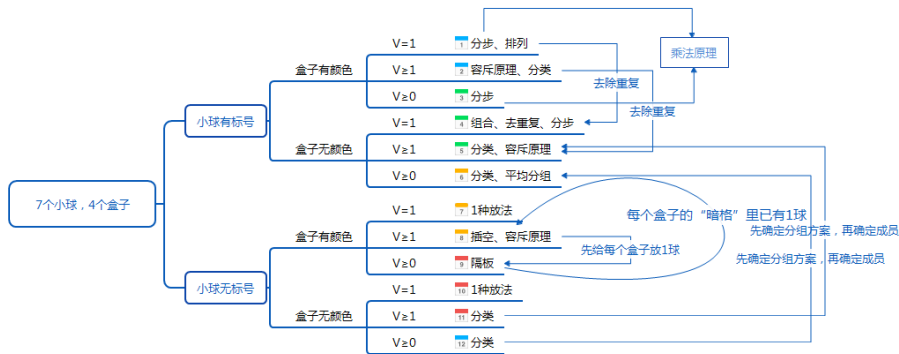


图: 计数问题结构框架与计数技巧方法思维导图

# 提升长时记忆效果的策略

## 构建自己的联系丰富的数学知识网络

自己构建体系的记忆效果远远优于别人告诉我的同样内容. 把记忆的东西与大脑中原有的东西联系起来, 这种联系的神经通路越多, 对记忆对象的描述就越全面, 记忆对象会更加鲜活.



# 提升长时记忆效果的策略

## 构建自己的联系丰富的数学知识网络

自己构建体系的记忆效果远远优于别人告诉我的同样内容. 把记忆的东西与大脑中原有的东西联系起来, 这种联系的神经通路越多, 对记忆对象的描述就越全面, 记忆对象会更加鲜活.

### 数学知识网络中的各种联系

- 可以加深我们对知识的理解, 从而使得减轻记忆负担, 巩固记忆持效性
- 有利于通性通法的提炼, 从而使得记忆提取线索更为精准, 便于存储信息的检索、回忆、再现
- 有利于联想、发散性思维和创造性思维





# 提升长时记忆效果的策略

对自传体记忆的研究表明自身知识网络的构建方式决定了记忆水平. 记忆存储的信息不止知识与事实, 还包含个人体验、感情色彩、情绪状态以及时空上的参照点等信息. 把知识绑定自己的积极体验, 主动建构知识网络, 使之成为自传体记忆的一部分

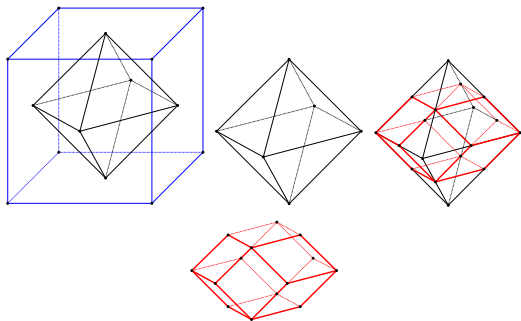
表 4 自传体记忆特征<sup>55</sup>

特征	记忆类型	
	自传体记忆	语义记忆
自我参照	高	极少
认知体验	一直出现	极少出现
个人解释	频繁出现	极少
真实性	有改动	表现出社会一致性
记忆持续期	几年或更长	几天到几年或更长
情景具体化的感觉与直觉	一直出现	从未出现
想象	频繁出现	少有出现



# 提升长时记忆效果的策略

## 视觉技能发掘长时记忆潜力

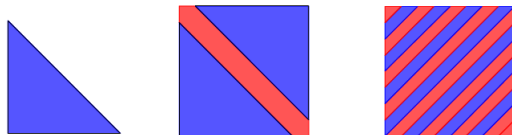


# 提升长时记忆效果的策略

## 令数学记忆生动起来的队列阵形变换

问题 1、求  $\sum_{0 \leq i < j \leq 50} C_{50}^i C_{50}^j$  除以 31 的余数. (南开大学 2008 年保送生考试试题)

$$\text{解: } \sum_{0 \leq i < j \leq 50} C_{50}^i C_{50}^j = \frac{1}{2} \left( \sum_{i=1}^{50} \sum_{j=1}^{50} C_{50}^i C_{50}^j - \sum_{i=1}^{50} C_{50}^i C_{50}^i \right) = \frac{1}{2} \left( \sum_{i=0}^{100} \sum_{j=0}^i C_{50}^i C_{50}^{i-j} - \sum_{i=0}^{50} C_{50}^i C_{50}^{50-i} \right)$$



三角阵型

→

方阵型

横行逐行相加

→

斜行逐行相加

图 43 队列阵形变化示意图一

$$= \frac{1}{2} \left( \sum_{i=0}^{100} C_{100}^i - C_{100}^{50} \right) = \frac{1}{2} (2^{100} - C_{100}^{50}) = 2^{99} - \frac{1}{2} C_{100}^{50} \equiv 16 \times (31+1)^{19} \equiv 16 \pmod{31}$$

# 发挥多重记忆系统合力的策略

## 巧设数学活动——寓教于乐中的内隐记忆

演绎推理的方式，记忆的对象不仅有公式，也有蕴含其中的数学思想，它的学习一般需要有意注意，相应的记忆也是外显记忆；而操作摸索的过程中更多的是无意注意，摸索得到的结论不经意间已留下深刻印象，相应的记忆是内隐记忆。内隐记忆存储的信息丰富生动完整，能保持的时间更长，在迁移提取时比外显记忆更稳定。

★ 两种方式协同进行时学习效果最好



# 发挥多重记忆系统合力的策略

## 巧设数学活动——寓教于乐中的内隐记忆

演绎推理的方式，记忆的对象不仅有公式，也有蕴含其中的数学思想，它的学习一般需要有意注意，相应的记忆也是外显记忆；而操作摸索的过程中更多的是无意注意，摸索得到的结论不经意间已留下深刻印象，相应的记忆是内隐记忆。内隐记忆存储的信息丰富生动完整，能保持的时间更长，在迁移提取时比外显记忆更稳定。

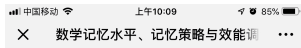
★ 两种方式协同进行时学习效果最好

## 规范流程与提炼靶事件编码对前瞻记忆的促进

- 规范严谨的程式让前瞻记忆成为无需意识的自动激活加工
- 提高靶事件的编码质量降低前瞻记忆对注意分配的要求



# 问卷



## 数学记忆水平、记忆策略与效能调查问卷

\* 10. 给您一个代数式:

$$x^3y + y^3z + z^3x + xy^2 + yz^2 + zx^2$$

与一个移动电话号码: 15971243275

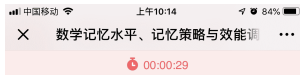
您觉得可以更快记住的是:

☐ 代数式

☐ 移动电话号码

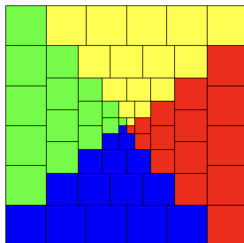
下一页

10/36页



## 数学记忆水平、记忆策略与效能调查问卷

\* 27. 短时间内 (不超过30秒) 您能记住下面这幅图吗?  
(您有45秒时间可以浏览这个页面)



☐ 肯定可以

☐ 能大致回忆出



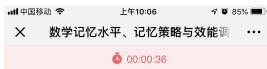
# 统计检验方法选择

## 数据统计分析——常用统计检验方法

2017年04月27日 15:16:47

		单变量	因变量			
			连续变量	有序分类变量	无序分类变量	二
单自变量	连续	样本t检验	相关分析 回归分析	有序分类的Logistic回归	无序分类的Logistic回归	二
	有序分类	单样本秩和检验	单因素方差分析，结果解释时利用有序信息	秩相关分析、CMH x2	可将自/因变量交换后分析	可
	无序分类	单样本x2检验	单因素方差分析	多样本秩和检验（H检验）	x2检验，深入分析可用对数线性模型	x2
	二分类	二项分布确切概率法	两样本t检验	两样本秩和检验（W检验）	x2检验	四
	连续变量	—	线性回归模型	有序分析的判别分析，有	判别分析、无序分类	判

# 短时记忆容量



华东师范大学  
EAST CHINA NORMAL UNIVERSITY

数学记忆水平、记忆策略与效能调查问卷

6. 如果您用较快的速度扫视下面这行数字与字符五遍

497&3251?2780@3S2!74

按从左往右顺序您最多可以记住：（您有45秒的时间可以浏览这个页面）

☐ 5个以下

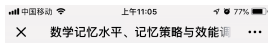
☐ 5到9个

☐ 9个以上

☐ 全部

< >

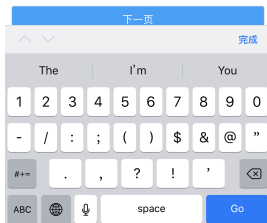
15 秒后才可以点下一页，  
45 秒后自动跳往下一页



华东师范大学  
EAST CHINA NORMAL UNIVERSITY

数学记忆水平、记忆策略与效能调查问卷

7. 请将您刚才记住的数字与字符按从左到右的顺序尽可能输入（请不要回看上一个页面喔）：



要输入的数字与字符在同一软键盘，减小操作影响

## 验证结论

短时记忆容量的平均值为：6.596 个字符





# 元记忆

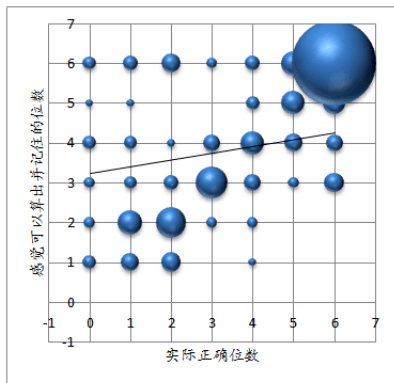


图 58 元记忆监控水平分布气泡图

表 7 元记忆水平与长时记忆水平斯皮尔曼二列等级相关分析表

## Correlations

			level of metamemory	level of long term memory
Spearman's rho	level of metamemory	Correlation Coefficient	1.000	.229**
		Sig. (2-tailed)		.005

# 长时工作记忆

Correlations

			working memory	level of education
Spearman's rho	working memory	Correlation Coefficient	1.000	.204**
		Sig. (2-tailed)	.	.008
		N	170	170
	level of education	Correlation Coefficient	.204**	1.000
		Sig. (2-tailed)	.008	.
		N	170	170

\*\* . Correlation is significant at the 0.01 level (2-tailed).

Correlations

			workin memor
Spearman's rho	working memory	Correlation Coefficient	
		Sig. (2-tailed)	
		N	170
	grade can teach	Correlation Coefficient	.554**
		Sig. (2-tailed)	.000

## 长时工作记忆相关调查结果

- “专家记忆优势效应”在数学领域同样存在
- 高效的提取可以使一部分长时记忆成为工作记忆的有效延伸



# 视觉技能、图象记忆、空间记忆

表 11 男女的心理旋转能力差异卡方检验表

sex \* level of mental rotation Crosstabulation

			level of mental rotation			Total
			0	1	2	
sex	female	Count	32	57	74	163
		% within sex	19.6%	35.0%	45.4%	100.0%
	male	Count	18	41	95	154
		% within sex	11.7%	26.6%	61.7%	100.0%
Total	Count		50	98	169	
	% within sex		15.8%	30.9%	53.3%	

## 心理旋转能力相关调查结果

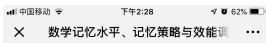
- 男女的心理旋转能力有显著性差异
- 能力结构差异使得记忆策略的使用效果因人而异, 对记忆策略也有不同偏好

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	8.893 <sup>a</sup>	2	.012
Likelihood Ratio	8.958	2	.011



# 视觉技能、图象记忆、空间记忆



## 数学记忆水平、记忆策略与效能调查问卷

- \* 18. 有人电话里告诉你一个  
明天要用到的查询码

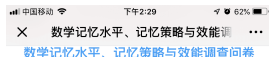
"1236 9874 5741 5963",

如果仅凭记忆您需要:

- ☐ 电话里记住即可
- ☐ 当即超饱和复述 (背熟了仍旧再复述几次)
- ☐ 间隔一段时间就复述
- ☐ 感觉办不到 (睡一觉会忘记)

< >

仅有 13 人 (约占总人数的 4.1%) 选择第一选项 (认为可以比较轻松的记住)



- \* 19. 有人用当下热门的  
手势密码的形式给你看了下

明天要用到的查询码

"1236 9874 5741 5963",



您能扫一眼就记住这个查询码吗?

☐ 可以

< >

有 302 位 (约占总人数的 95.3%) 选择第一选项 (认为可以比较轻松的记住)

## 结论

合适的记忆策略可以  
大幅提高记忆效能,  
视觉技能在记忆中发  
挥着重要的作用。



# 多重记忆系统

## 内隐记忆

调查证实记忆确实以多重系统的方式存在，即有需要意识努力的外显记忆，也存在无意间的内隐记忆



# 多重记忆系统

## 内隐记忆

调查证实记忆确实以多重系统的方式存在，即有需要意识努力的外显记忆，也存在无意识的内隐记忆

## 闪光灯记忆

调查证实即使关于数学大部分人也存在过闪光灯记忆，这种方式不仅记忆负担轻而且持效性好



# 多重记忆系统

## 内隐记忆

调查证实记忆确实以多重系统的方式存在，即有需要意识努力的外显记忆，也存在无意间的内隐记忆

## 闪光灯记忆

调查证实即使关于数学大部分人也存在过闪光灯记忆，这种方式不仅记忆负担轻而且持效性好

## 前瞻记忆

调查证实确有前瞻性记忆水平偏低致使“查漏补缺”、“错题本”也作用不大的情况



# 建议

- ❶ 避免教授碎片化的数学知识
- ❷ 注意培养借助符号沟通能力，提高抽象思维能力
- ❸ 正视学生个性化差异，呈现多样化记忆策略
- ❹ 教学设计兼顾外显记忆与内隐记忆
- ❺ 深度加工记忆信息、优化编码、尝试多样化提取
- ❻ (热爱运动、科学作息、阳光生活)





# 研究中的不足与发展

- 知识结构** 数学教育心理学补短板、科普知识更新、掌握提升效率的教育新技术手段
- 行动研究** 不够深入，将继续“计划-行动-观察-反思-修订计划-行动-观察-反思……”的深入探索，进一步尝试教育实验研究的对比验证
- 调查研究** 变量控制不够严格，样本范围未能做到精准锁定，时限、关联问题设置比较粗放



# 感谢华南师范大学为我们所创设的系统学习机会



图: 华东师范大学 2014、2015 级单证在职数学教育硕士 (学科教学) 合影



# 感谢各位专家的指导！

指导老师

李建国



# 感谢各位专家的指导！

指导老师

李建国

答辩人

付海林

2016022292

软件工程（学术型）

2016 级



# 感谢各位专家的指导！

指导老师

李建国

答辩人

付海林

2016022292

软件工程（学术型）

2016 级

