

代数应用题项目生成中的认知过程与任务特征分析*

杨向东

(华东师范大学课程与教学研究所, 上海 200062)

摘要 认知项目设计起源于理论驱动的测量模式,要求在详尽阐述测量建构理论的基础上,鉴别测验项目类型的任务特征,建立和标定与测量建构认知变量之间的对应关系,从而通过系统变换任务特征生成项目。本文以 Greeno (1994)提出的代理者-情境交互作用(agent-situation interaction)的问题解决观为基础,对代数应用题解决的认知过程和任务蕴含性(affordance)进行了系统分析。具体而言,代数应用题被视为以文本形式呈现的,嵌套在现实情境中的某种量化结构。在代数应用题解决的转译、整合、计划、执行的认知过程中需要形成三种不同的表征,即文本基、情境模型和问题模型。其形成及解决过程受到个体问题图式知识的深刻影响。对应于不同表征方式,对代数应用题语法/语义结构、量化元素及结构、问题情境特征进行了系统分析和综合。该分析表明,为了保证代数应用题项目生成的科学性,当前研究的关键在于建立任务特征与问题解决认知变量之间的经验关系,并在问题领域整体水平上检验这种关系的可推广性。

关键词 理论驱动的测量模式;项目生成;代数应用题;认知过程;任务特征

分类号 B841

1 引言

自动化项目生成是近年来心理与教育测量中比较活跃的领域(Alves, Gierl, & Lai, 2010; Daniel & Embretson 2010; Gierl & Haladyna, 2012)。自动化项目生成的基础是测验项目的认知设计技术(Embretson, 1998)。该技术要求对测验任务类型进行详尽的任务分析和认知研究,明确项目解决所需认知过程、策略和知识结构,开发形成相应的项目认知模型。在此基础上,鉴别测验项目类型的系列任务特征,并建立任务特征与项目解决认知变量间的对应关系,标定不同特征对项目反应机制的影响力度。这样,项目设计便可以通过系统变换项目类型中的任务特征,操纵项目解决过程中涉及到的认知成分及其负荷,生成具有指定测量学特征的项目(Embretson & Yang, 2007)。

测验项目的认知设计技术起源于理论驱动的

测量模式(Bouwmeester & Sijtsma, 2004; Embretson, 1983, 2004; Lohman & Ippel, 1993)。理论驱动的测量学研究范式旨在变革传统测量学去理论化、功能性和实用性取向的特征。该范式强调通过理论分析和实证研究,建立测验任务反应机制与测验建构内在表征之间的对应关系(Irvine & Kyllonen, 2002),使测量建构的认知理论成为建构界定、项目设计、结果评分、建模分析、结果解释和运用等环节的驱动原则(Embretson & Gorin, 2001),从而确保对个体进行推断和解释的效度,以及后继决策或干预措施的成效。

近年来,测验项目的认知设计法被运用到心理旋转、隐藏图形识别、瑞文矩阵推理、空间折叠、类比推理、序列完成等传统智力和能力倾向测验的任务类型(Irvine & Kyllonen, 2002)。随着研究的深入,对阅读理解、数学问题解决等复杂认知领域问题的研究也正在逐步展开(Arendasy & Sommer, 2007; Daniel & Embretson 2010; Gorin, 2005)。本研究正是在测验项目认知设计的视角下,对代数应用题的认知加工过程以及问题结构进行分析,鉴别代数应用题的任务特征,对后继项目生成研究奠定基础。

收稿日期: 2012-04-25

* 国家自然科学基金(31171000)和上海市教育科学2011年度重点项目(A1117)资助。

通讯作者: 杨向东, E-mail: xdyang50@hotmail.com

2 代数应用题研究的意义和模式

代数应用题属于具体领域的复杂认知任务。问题解决一直是心理学研究的重要领域(Eysenck & Keane, 2000)。早期研究多采用知识贫乏任务,如“九点阵”问题(nine-dot problem),河内塔问题(“Tower of Hanoi” problem)等。随着研究深入,问题解决研究逐渐转移到数学、科学、工程等知识丰富领域的复杂认知任务(Anderson, 1993; Snow & Lohman, 1989; Lohman, 2000)。具体领域任务的研究有助于重要理论问题的提出。早期研究的任务特征决定了这些问题不会被涉及到(Mayer, Larkin, & Kadane, 1984)。具体领域问题的任务解决过程涉及到个体认知结构与领域知识的整合,问题解决策略从通用的、领域一般性的启发式方法如手段-目的分析法(means-end analysis)开始向领域相关性策略(domain-dependent strategy)或领域特异性策略(domain-specific strategy)转变。在这个意义上,代数应用题研究对认知研究具有重要的理论意义。

此外,代数应用题还一直是心理或教育量表中的重要问题类型之一。许多著名的智力和能力倾向量表如斯坦福-比纳量表(SB5; Roid, 2003)、韦克斯勒成人智力量表修改版(WAIS-R; Wechsler, 1997)和儿童智力量表第三版(WISC-III; Wechsler, 1991),美国军队职业能力倾向量表(ASVAB; Ree & Carretta, 1995)都运用算术或代数应用题来测量个体的量化推理能力(quantitative reasoning ability)。在晶体-流体智力理论(Horn & Cattell, 1966)和三层结构智力理论(three-stratum-theory of intelligence; Carroll, 1993)中,量化推理能力是智力结构中一个非常稳定和重要的二阶因子。因此,代数应用题研究还是测量人类学习、思维和智能的重要领域。

然而,在传统测量学研究范式下,代数应用题测量学研究与认知研究是彼此分离的。严格来讲,代数应用题解决过程是个体数学认知结构与特定问题情境之间交互作用的动态过程(Sebrechts, Enright, Bennett, & Martin, 1996)。因此,问题解决研究应该既关注问题解决者的认知因素,也考虑问题情境特征。认知研究通常关注个体解决不同数学问题的共同模式、规律或认知机制。虽然认知研究也考虑问题情境特征或结构的变动,但出

发点是检验数学问题解决认知理论的依存条件或推广程度,并非关注实验材料本身。而对心理测量学者来讲,问题情境或结构特征、变动规律和呈现顺序成为了关注的焦点。明确不同问题特征的蕴含性(affordance),即解决具有特定特征的问题可能涉及到的知识基础、认知负荷、认知或元认知策略等,是测验项目设计和测验编制的核心问题。对测验任务反应机制的研究则是联接这两种不同研究模式的桥梁。通过对问题认知加工过程和任务特征的详尽分析,指向项目生成的代数应用题研究旨在建立任务特征与问题加工过程之间的关系,从而可以整合两种不同的研究模式(Embretson, 1998)。它不仅是自动化项目生成和认知诊断的理论基础,还对相关认知理论的构建产生深远的影响。通常,旨在构建理论的实验研究是在有限的情境、样本和任务类型下展开的,缺乏整合不同实验结果的系统方法(Cronbach, 1957)。因此,不同研究发现的不一致有可能仅仅是因为实验条件或任务特征不同而已,而非实质意义上的理论分歧。如果以代数应用题解决的认知理论为基础,建立该测量建构在整个任务领域范围的解释性测量尺度(explanatory scale),就可以以此为理论框架,整合和标定不同实验研究的发现,从而使理论驱动测量成为量化推理能力的理论构建和推广不可或缺的手段。

3 代数应用题的认知加工过程

代数应用题解决通常被认为是一个多阶段的复杂认知活动。比如,假设向适龄个体呈现下面这个问题:

案例 1: 小丁丁步行去少年宫。他平均每分钟走 75 米。小丁丁走了 8 分钟后,爸爸骑车追赶,他每分钟比小丁丁多走 120 米。爸爸几分钟后在途中追上小丁丁?

要正确解决这个问题,个体首先需要阅读问题陈述,正确理解其中的关键名词或术语,以及该陈述所要表达的含义,形成合理的内在表征。比如,个体需要能够根据上下文,知道第二句中的“他”是指小丁丁,而第五句中的“他”是指爸爸。另外,个体需要能够根据问题陈述辨析和形成问题中的关键变量及其取值。如根据“他平均每分钟走 75 米”,个体需要理解“小丁丁的步行速度”是一个关键变量,该变量的取值是“75 米/分钟”。该

阶段通常称为问题转译(problem translation)阶段(Mayer et al., 1984)。问题陈述的解码过程被认为和文本理解过程在原则上是一致的(Kintsch & Greeno, 1985)。经过问题转换, 问题陈述被表征成为一系列的命题(proposition), 又称文本基(text base; Kintsch, 1998)。

第二个阶段是个体根据问题陈述中所表述或蕴含的情境或数量关系, 形成对问题结构的整体表征, 即问题整合(problem integration)阶段(Mayer et al., 1984)。需要指出的是, 个体这种整体认识未必一定是量化的, 也有可能是对问题情境特征和结构的质性模型, 即所谓的情境模型(situational model; Kintsch & Greeno, 1985; Nathan, Kintsch, & Young, 1992)。例如, 虽然在案例1中没有明确表明, 个体需要能够判断爸爸追赶路线和小丁丁步行路线是相同的。同样的, 个体需要能够推测出小丁丁和爸爸相遇时所走的路程是相同的。这些基于生活常识的理解和推测使得个体能够对问题中的各种信息(人物、行动、变量等)的语义和时空关系进行整合, 形成对问题情境的整体认识。这种认识未必是变量之间准确的数量关系, 有可能只是一种质性的变量关系推断。

然而, 虽然有可能存在质性的情境模型, 个体要解决上述问题, 必须能够形成对各种变量之间的数量关系, 即所谓的问题模型(problem model; Kintsch & Greeno, 1985; Nathan et al., 1992)。问题模型是个体形成问题解决的算术式或代数式的基础。在已有研究中, 并没有明确的证据表明质性的情境模型一定先于量化的问题模型形成(Koedinger & Nathan, 2004)。但是问题模型的形成必须能够将情境模型中的各种限制考虑在内。后者基于个体一般性的生活常识, 前者则建立在个体的算术或代数知识基础之上。合理问题模型的形成在于个体能否建立两者之间的协调, 能够构建满足变量之间关系和具体情境限制的算术或代数式。

问题转译和整合阶段构成了个体对问题的理解过程, 即传统意义上的审题阶段。之后, 个体进入问题解决阶段。通常, 个体根据审题阶段的问题表征制定问题解决计划(solution planning)。不同的问题表征有可能导致个体采用不同的问题解决策略。例如, 在案例1中, 如果个体关注爸爸追上小丁丁时两人所走路程相同, 相应的代数式可

能为 $75 \times 8 + 75t = (75 + 120)t$, 其中 t 代表爸爸追上小丁丁时所用时间。如果个体意识到爸爸追上小丁丁时所花的时间就是爸爸弥补小丁丁先走距离所用的时间, 那么相应的代数式则变成 $75 \times 8 = 120t$ 。在确定解题计划和策略后, 个体应用所学数学或者代数知识执行计划(solution execution)。对某些个体而言, 监测自己的运算过程以及比较运算结果与问题模型之间的一致性也是执行计划的重要环节(Mayer, 1998)。

问题转译、整合、计划、执行构成了数学应用题解决的一般环节。然而, 不同数学专长的个体并非经历相同的认知过程(Hegarty, Mayer, & Monk, 1995)。对拥有大量问题图式(schema)的个体而言, 问题的表征过程更像是一个假设-检验(hypothesizing-and-testing)的演绎推理过程(Chi, Feltovich, & Glaser, 1981)。问题图式是指个体所拥有的围绕特定问题类型而组织起来的各种知识(Mayer et al., 1984)。良好的问题图式既包括该问题类型中常见的变量类型和问题结构, 也包括这些结构所处的典型问题情境的信息(Arendasy & Sommer, 2007), 以及与问题解决相关的策略性、陈述性或程序性知识(Marshall, 1995)。对这些个体而言, 问题的头几句陈述, 甚至头几个词语就可能激活他们长时记忆中储存的相关问题图式, 从而迅速形成对问题类型的暂时性假设(Hinsley, Hayes, & Simon, 1977)。例如, 在案例1中, 个体读完第一句陈述, 就可能形成该问题属于“行程”问题的预期。第二句陈述则会强化这种预期。此时, 行程问题中的“追及”、“相遇”或“往返”等问题图式都有可能成为备择假设。当个体读到“小丁丁走了8分钟后, 爸爸骑车追赶”时, “行程追及”便成为了最为可能的假设, 引导个体进一步确认这种预期。因此, 问题图式引导问题解决者 with 问题陈述之间的“对话”过程, 使得问题转译过程变成了预期-验证式的模式匹配过程。同时, 图式提供了个体整合情境模型和问题模型的模板, 从而极大降低了问题理解的认知负荷(Sweller, 1989)。问题图式还有助于激活个体长时记忆中与该问题类型相关的问题解决策略、公式、定理等知识, 协助个体列出问题等式, 选择合理策略并正确解题。大量研究表明, 拥有大量的领域知识和问题图式是领域专家的关键特征之一(Anderson, 1993; Feldon, 2007)。

当个体没有形成或拥有很少问题图式时,绝大多数问题都是陌生问题。这种情况下,问题转译过程主要是一种文本驱动的自下而上的解码过程(De Corte, Verschaffel, & De Winn, 1985)。个体主要依赖其语言知识和世界知识逐句逐句的理解问题陈述。然后分析不同变量的类型及其相互关系,并综合问题情境的限制,才能形成整合的问题表征。此时,问题理解过程往往具有很高的认知负荷(Sweller, 1989)。个体通常借助于手段-目的法(means-end method)来建立各种子目标,或者通过借助于画图或列出每个陈述对应算式的外部表征方法来降低认知负荷。但即便如此,个体由于缺乏足够的认知资源,往往会出现各种各样的表征或运算错误。此外,研究表明,如果个体只是形成了情境模型,没有形成相应的问题模型时,则有可能不用列算式或代数式的策略解题,而是运用倒推(unwind)、比例设置(ratio setup)、猜测检验(guess-and-test)、基于情境模型的模拟(model-based simulation)等非正式策略解题(Koedinger & Nathan, 2004; Sebrechts et al., 1996)。非正式问题解决策略的激活通常受问题某些具体特征的影响,往往只适用于某一类或几类问题,缺乏普遍性。比如,Sebrechts 等人(1996)发现,当问题解决需要多次涉及到某个未知变量(见下面的双重引用问题),或者需要设定多个未知变量时,个体可能采用基于情境模型的模拟策略,而不是通过代数式进行解答。与列代数式策略相比,模拟策略既低效又缺乏准确性。

4 代数应用题的任务特征分析

代数应用题解决的认知过程受题目具体任务特征及其变动的影响。常见的代数应用题可以理解为是以文本形式呈现的,嵌套在特定现实情境中(real-life situation)的各种变量之间的某种量化结构(quantitative structure)。因此,可以从三个方面分析代数应用题的复杂程度:(1)应用题的语言陈述方式,(2)相关变量的类型、数量和彼此关系,(3)问题所处的具体情境。

4.1 应用题的语言陈述方式

如果问题陈述的理解过程符合文本理解的基本原则,可以预期问题陈述的语言特征应该与问题转译的难易程度有关。早期研究的确发现,个体代数应用题的解题能力与其语言能力之间存在

一定程度的相关(两者一阶相关系数从 0.36 到 0.87 不等;Aiken, 1971)。新近研究发现,简化数学问题陈述的某些语言特征,如降低字词总数、减少生僻或艰涩词汇、避免模棱两可的词或多义词、将被动时态改为主动时态、将复杂句改成简单句等,的确会对题目难度产生一定影响(Abedi, 2004; Shaftel, Belton-Kocher, Glasnapp, & Poggio, 2006)。然而,这种改变通常只对语言水平较低的被试有所帮助,而且影响力度通常很小(Abedi & Lord, 2001; Johnson & Monroe, 2004)。Shaftel 等人(2006)研究发现,简化问题陈述的语言特征在个体成绩变异中所能解释的方差比例非常小(以调整后的 R^2 为指标,在 4、7 和 10 年级中分别为 0.134、0.07 和 0.041)。

与表面语言特征相比,问题转译难度可能更多与问题陈述的结构特征(structural feature)有关。结构特征主要包括问题陈述的语法结构(syntactic structure)和语义结构(semantic structure)(Loftus & Suppes, 1972)。Mayer 等人(1984)区分了数学问题陈述中三种不同的语法结构:赋值(assignment)、关系(relation)和问题(question)。赋值陈述给某个变量赋值,例如“他平均每分钟走 75 米。”关系陈述表达了两个变量之间的数量关系,例如“爸爸每分钟比小丁多走 120 米。”问题陈述询问未知变量的取值,例如“爸爸几分钟后追上小丁?”研究表明,不管是成人还是儿童,在理解关系陈述时都会出现更多的困难(Lewis, 1989; Lewis & Mayer, 1987; Mayer, 1982)。

问题陈述的语义结构也会影响问题的表征过程。Lewis 和 Mayer (1987)在研究不同类型的比较应用题(compare problem)(见表 1)时发现,当关系陈述所包含的未知量是该陈述的宾语,同时关系术语(如比……少)与所要求的数学算法(如相加)相冲突时,个体容易出现转译错误,并采用与题意相反的算法;而当关系陈述所包含的未知量是该陈述的主语,同时关系术语(如比……多)与所要求的数学算法(如相加)相一致时,转译错误很少出现。这两类问题分别被称为语言不一致性问题(inconsistent language problem)和语言一致性问题(consistent language problem)。他们认为,这种现象表明个体在问题表征过程中对信息呈现顺序存在某种偏好。当问题陈述的语义结构与个体信息呈现偏好不符时,个体需要将呈现的语义结

表 1 两类不同的比较应用题(改编自 Lewis & Mayer, 1987)

运算	问题类型	
	语言一致	语言不一致
加法	约翰有 3 个石子, 汤姆比约翰多 5 个石子, 汤姆有几个石子?	约翰有 3 个石子, 他比汤姆少 5 个石子, 汤姆有几个石子?
减法	约翰有 8 个石子, 汤姆比约翰少 5 个石子, 汤姆有几个石子?	约翰有 8 个石子, 他比汤姆多 5 个石子, 汤姆有几个石子?

转换成偏好的语义结构, 从而造成问题理解和表征中容易出错。此外, 当关系术语为标示性(marked)时, 语义结构偏好对问题表征的影响力度加大。按照 Clark (1969)的词法标示理论(theory of lexical marking), 比……少, 比……短, 比……年轻等关系术语为标示性的, 而比……多, 比……高, 比……老等为非标示性的(unmarked)。非标示性术语如比……多和它的对应标示性术语如比……少是同一测量尺度上的不同参照点。基于这一理论, Kintsch 和 Lewis (1993)认为, 比……多和比……少对应于同一心理数线(mental number line)。由于心理数线的形成与非标示性术语是一致的(即都是从小到大), 而标示性术语则存在一个潜在的限制, 因此, 个体在整合含有标示性术语的问题时所经历的认知过程, 和整合非标示性术语的问题存在质的区别。这一区别导致了含有不同关系术语问题在表征难度上的差异。

4.2 应用题中的变量数量、类型及其关系

在问题转译基础上, 形成对问题的整体表征是代数应用题解决的关键环节。该环节除了要求个体具备必要的数学概念和原理外, 主要取决于问题对个体所施加的注意资源(attention resources)和工作记忆容量(working memory capacity)的需求(Sweller, 1989)。从任务特征的角度来看, 问题解决认知需求主要受到两个方面的影响: (1)解题过程中需要考虑的数据数量及其复杂性, 和(2)获取最终答案所需的中介结果数量及其复杂性(Stillman, 1996)。这些因素都与代数应用题中变量数量、类型及其关系有着紧密联系。

变量数量和类型 变量(variable)数量在一定程度上标示了应用题问题模型的复杂程度。一般情况下, 变量越多, 问题整合时涉及到的信息量越大, 对个体所施加的认知负荷越大。例如, 案例 1 涉及到 6 个变量和 3 个已知数值, 其中 6 个变量

分别为丁丁的速度、爸爸的速度、丁丁所用时间、爸爸所用时间、丁丁所走距离和爸爸所走距离。三个数值为 75 米, 8 分钟和 120 米。相比之下, 表 1 中的比较问题(以加法运算的一致性问题为例)包括了 2 个变量, 即汤姆的石子数和约翰的石子数, 和 2 个已知数值, 即 3 个石子和 5 个石子。单从变量数量来看, 案例 1 要比表 1 的比较问题复杂。Lane (1991)证实了这一点。她通过变换问题陈述数量来改变问题的元素数量(变量、数值、变量关系), 发现含有较多元素的问题难度更大。

除数量之外, 变量类型和取值都会对问题解决过程产生影响。Shalin 和 Bee (1985)区分了代数应用题中两种不同类型的变量, 广延变量(extensive variable)和集中变量(intensive variable)。广延变量是指自然界中事物的基本属性(primary property), 如长度、距离、时间、体积等等, 而集中变量是指事物的二阶属性(secondary property), 反映了事物两个基本属性之间的某种映射关系(mapping between two extensives), 比如速度是单位时间内经过的距离, 类似的包括密度、温度等等。研究表明, 集中变量会导致问题的难度增加(Nesher & HersHKovitz, 1994)。此外, 变量取值为整数时, 问题相对简单, 而取值为分数、小数时则相对复杂(Koedinger & Nathan, 2004)。变量取值所引起的复杂性不仅仅局限于运算方面, 还包括因变量取值而导致的问题转译和整合阶段的困难。例如, 虽然案例 2 和 3 问题结构相同, 但是变量取值从整数到百分数的转变, 会导致个体在选择适合案例 3 的逆运算时遇到更大的困难(Tabachneck, Koedinger, & Nathan, 1995)。

案例 2: 某数的 15 倍是 100, 这个数是多少?

案例 3: 某数的 80%是 100, 这个数是多少?

类似的, 当变量取值从具体的数值转变成抽象的符号时, 如用 x 米/分钟取代 75 米/分钟, 个体需要能够在抽象水平上对问题的不同元素进行

整合,问题难度也会增加(Enright, Morley, & Sheehan, 2002)。这种效应在问题最终答案需要用符号表达式表示时更加明显(Sebrechts et al., 1996)。

问题结构和变量关系 问题结构是指问题中变量的某种特定组合模式。变量数量相同的代数应用题,复杂程度会因问题结构不同而迥然相异。影响问题解决过程认知负荷的众多因素,如子目标多少、信息转换或整合数量和难度、限制条件数量、运算次数和复杂性等,都与问题结构有着内在联系。

代数应用题结构的基本元素是变量(variable)和数值(quantity)。数值可能是某个变量的取值,如“(小丁丁)平均每分钟走 75 米”,也可能是两个变量之间的数量关系,如“(爸爸)每分钟比小丁丁多走 120 米”。问题结构中变量关系可以是两个变量之间的某种数量关系,如“爸爸的速度=丁丁的速度+120 米”,也可以是三个变量之间的某种数量关系,如“爸爸所走距离=爸爸的速度×爸爸所走时间”。不管哪种情况,问题结构的基本单位都可以界定为是由三个元素构成的一个三阶组合(triad),如“爸爸的速度”、“丁丁的速度”和“120 米”就构成了一个三阶组合。不同类型的三阶组合就构成了问题模型的各种基本结构(elementary structure; Sebrechts et al., 1996)。

复杂代数应用题的问题结构则是由不同的基本结构组合而成的。不同的基本结构可以有多种组合方式。对于包含有两个基本结构的代数学应用题,Shalin 和 Bee (1985)区分了 3 种基本结构的组合方式:层级或嵌套型(hierarchical/nested)、共享整体型(shared whole)和共享部分型(shared part)。案例 4、5 和 6 分别给出了三种组合方式的实际案例。案例 4 由两个三阶组合构成,即“爸爸年龄=6×小军年龄”和“爷爷年龄=爸爸年龄+30 岁”。第一个三阶组合中的总数“爸爸年龄”成为第二个三阶组合中“爷爷年龄”的一个组成部分,从而使得两个三阶组合形成嵌套关系。案例 5 中的两个三阶组合分别为“水蜜桃个数=12 只/箱×24 箱”和“水蜜桃个数=已卖 189 个水蜜桃+剩余水蜜桃个数”。两个三阶组合具有共同的上位总数“水蜜桃个数”,形成共享整体型结构。案例 6 中的两个三阶组合分别为“48 支红玫瑰=6 支红玫瑰/束×卖出的玫瑰花束”和“卖出的白玫瑰数=4 支白玫

瑰/束×卖出的玫瑰花束”。两个三阶组合具有共同的下位成分“卖出的玫瑰花束”,形成共享部分型结构。

案例 4: 小军今年 6 岁。爸爸的年龄是小军的 6 倍。爷爷年龄比爸爸大 30 岁,爷爷今年多少岁?

案例 5: 水果店运来 24 箱水蜜桃,每箱 12 只,已经卖去 189 只,还剩多少只?

案例 6: 花店的每束玫瑰花由 6 支红玫瑰和 4 支白玫瑰组成。中午时小明清点已经卖出的玫瑰花束,发现已经卖出了 48 支红玫瑰,白玫瑰卖出了多少支?

如果把上述三种组合方式称之为三种高阶结构(higher-order structure),复杂代数应用题的问题结构就是由不同数量的高阶结构组合而成。图 1 给出了案例 1 的问题结构。根据问题陈述,可以转译两个三阶组合,即“爸爸的速度=丁丁的速度+120 米”和“丁丁所用时间=爸爸所用时间+8 分钟”。根据问题情境和已有知识,还可以推断出三个三阶组合,即“爸爸所走距离=爸爸的速度×爸爸所用时间”、“丁丁所走距离=丁丁的速度×丁丁所用时间”、和“爸爸所走距离=丁丁所走距离”¹。因此,案例 1 可以认为包含了 5 个三阶组合。其中,根据“爸爸所走距离=丁丁所走距离”这一条件,可以将“丁丁所走距离=丁丁的速度×丁丁所用时间”转换为“爸爸所走距离=丁丁的速度×丁丁所用时间”,进而明确该三阶组合和“爸爸所走距离=爸爸的速度×爸爸所用时间”的关系为共享整体型结构。图 1 表明,案例 1 的问题结构总共包括了 2 个嵌套型、2 个共享部分型和 1 个共享整体型的高阶结构。

已有研究表明,含有两个三阶组合的代数应用题中,三种高阶结构是影响问题难度的一个重要因素(Shalin & Bee, 1985; Nesher & Hershkovitz, 1994)。这表明问题结构和变量关系分析提供了描述代数应用题复杂性的一种方法。然而,对于更为复杂的代数应用题,如何根据结构特征抽取能够表示问题复杂性的关键变量仍然是一个有待研究的问题。综观已有研究,可以发现几个刻画问

¹两个变量取值相等是两个变量之间数量关系的特例。这种情况仍然可以看作为三阶组合,即“爸爸所走距离=丁丁所走距离+0 米”。

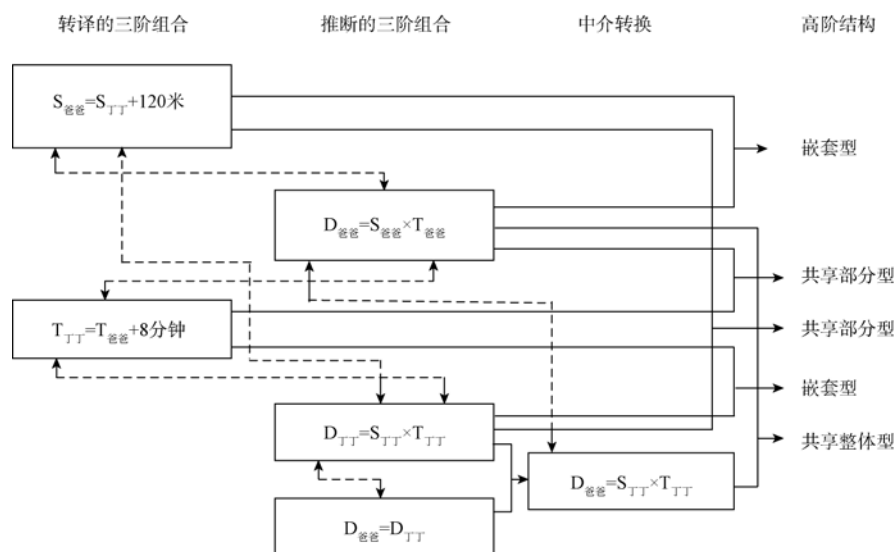


图 1 案例 1 中问题结构的基本和高阶结构

注：图中符号分别表示为： $S_{爸爸}$ =爸爸的速度， $S_{丁丁}$ =丁丁的速度， $T_{爸爸}$ =爸爸所用时间， $T_{丁丁}$ =丁丁所用时间， $D_{爸爸}$ =爸爸所走距离， $D_{丁丁}$ =丁丁所走距离。其中用粗体表示的 $T_{爸爸}$ 为所求变量。

题复杂性的常用变量。一个变量是问题中包含的分步等式(代数式)数量(number of partial equation; Arendasy & Sommer, 2007)。例如，案例 1 中“爸爸的速度=丁丁的速度+120 米”就是一个分步等式。根据图 1 可以看出案例 1 包含了 5 个分步等式。显然，分步等式对应于问题中的三阶组合。因此，该变量揭示了问题中基本结构的数量。问题包含的分步等式越多，问题模型的规模越大，个体在形成问题表征时需要整合的因素也就越多，在一般情况下问题也就越复杂。另外，每个分步等式界定了问题变量之间的一个限制条件。因此，分步等式数量同时也表示了问题解决所需满足的限制数量(number of constraints to be satisfied; Enright, Morley, & Sheehan, 2002)。从抽取变量的角度，这两个变量应该是等价的。

与分步等式数量密切相关的另一个变量是问题解决中子目标的数量(number of subgoal; Embretson & Daniel, 2008)。子目标定义了问题模型中变量间的嵌套层次(level of nesting; Sebrechts et al., 1996)。按照问题空间理论(Newell & Simon, 1972)，问题解决过程是个体通过各种认知算子(cognitive operator)实现从问题初始状态到目标状态转换的过程。在每个环节，个体选择认知算子的基本原则是尽可能缩小当前问题状态和目标状

态的区别。当无法直接实现转换时，一种常用策略是创建一系列彼此相连的子目标，通过逐步实现子目标来实现最终的转换(Anderson, 1993)。例如，案例 4 的目标状态是计算爷爷的年龄，然而根据问题初始条件无法直接求取。个体需要根据所给条件，先求爸爸的年龄，再求爷爷年龄。爸爸年龄因而成为计算爷爷年龄的子目标。子目标数量在一定程度上定义了问题空间的大小(Newell & Simon, 1972)，与个体问题表征和解决过程中需要整合的元素数量、转换程度、运算步骤等有着深刻的联系。研究表明，子目标数量是预测问题难度的重要因素(Daniel & Embretson, 2010; Embretson & Daniel, 2008)。

分步等式或子目标数量虽然较好的刻画了问题模型的整体规模，但是在刻画问题内在复杂性方面有所欠缺。比如，案例 4 和案例 7a 都包含有两个分步等式和两个子目标。然而，两个问题的内在复杂性是不同的。

案例 7a 小青和妈妈的年龄和是 45 岁，妈妈的年龄是小青年龄的 4 倍，问妈妈的年龄是多少岁？

图 2 给出了这两个案例的问题结构。对案例 4 而言，两个三阶组合中爸爸的年龄($A_{爸爸}$)为共享变量，构成嵌套型结构。根据已知条件“ $A_{小青} = 6$ 岁”和

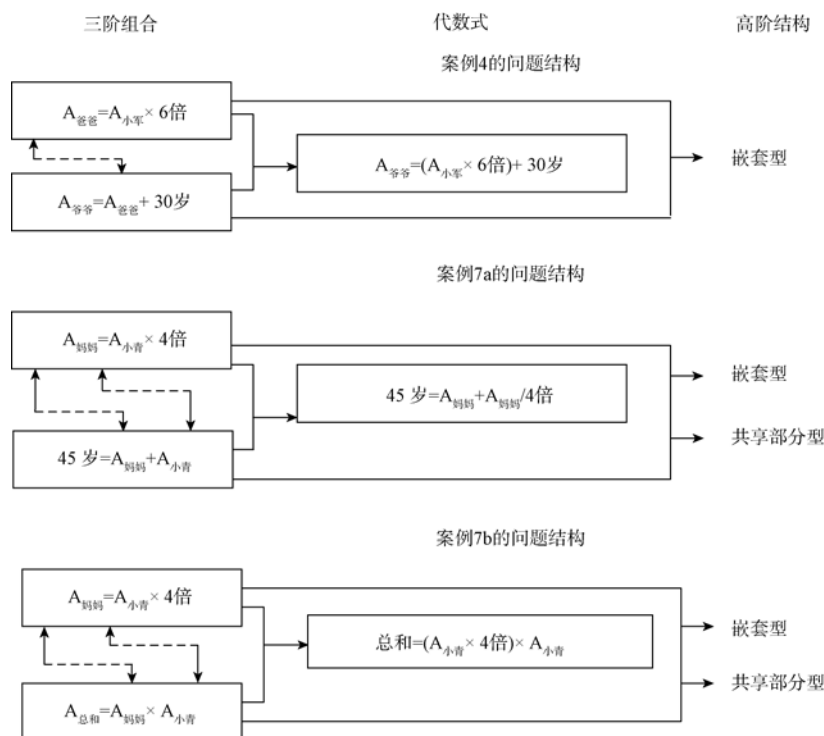


图2 案例4、案例7a和案例7b的问题结构

注：图中符号分别表示为： $A_{爸爸}$ = 爸爸的年龄， $A_{小军}$ = 小军的年龄， $A_{爷爷}$ = 爷爷的年龄， $A_{妈妈}$ = 妈妈的年龄， $A_{小青}$ = 小青的年龄。其中用粗体表示的变量为所求变量。

第一个三阶组合，即可求得子目标爸爸的年龄，进而通过两个三阶组合的嵌套关系求取爷爷的年龄。相比之下，案例7a的两个三阶组合中妈妈和小青的年龄($A_{妈妈}$, $A_{小青}$)都为共享变量。以妈妈的年龄为中介变量，两个三阶组合构成嵌套型结构。同时，以小青的年龄为中介变量，两个三阶组合又构成共享部分型结构。与案例4不同，根据已知条件，案例7a的两个三阶组合在单独的情况下都无法求解。只有在将两个三阶组合整合成一个代数式才有可能。这表明案例7a中变量之间的关联程度更高。按照认知负荷理论(Paas, Renkl, & Sweller, 2003; Sweller, van Merriënboer & Paas, 1998), 不同问题元素之间的互动程度(element interactivity)是造成个体问题解决认知负荷的内在原因。当元素之间关联程度较少时，个体可以依次理解或解决每个元素，无需同时考虑其他元素。而对于元素互动程度很多的问题，只有当所有的元素及其相互关系同时被加工时，才能真正理解问题本身。显然案例7a中变量之间互动程度

更多。因而同样情况下，该问题所造成的认知负荷更高。Koedinger, Abilali 和 Nathan (2008)将这两类问题分别称为单一引用问题(single-referenced problem)和双重引用(double-referenced problem)。在单一引用问题中，未知量在问题模型中出现或引用了一次(见图2中案例4的代数式)，而在双重引用问题中，未知量出现或引用了两次(见图2中案例7a的代数式)。因此，表征这两类问题的一个变量是问题解决代数式中未知变量出现的次数(number of times unknown variable occurs; Sebrechts et al., 1996)。双重或多重引用问题要求个体能够用某个变量的抽象符号来表征其他相关变量或答案。因而，这类问题通常难度较大(Embretson & Daniel, 2008)。

未知变量位置 需要指出的是，问题解决代数式中未知变量出现的次数与问题结构之间并不是一一对应的关系。例如，案例7b与案例7a具有相同的问题结构(见图2)，但案例7b中，根据已知条件“ $A_{小青}=9$ 岁”和第一个三阶组合，即可求得

子目标妈妈的年龄,继而求取未知变量的值,而无需像案例 7a 那样必须整合两个三阶组合才能求解。相应的,案例 7b 的代数式中未知变量仅仅出现 1 次,而非 2 次。

案例 7b 小青今年 9 岁,妈妈的年龄是小青年龄的 4 倍,问妈妈和小青的年龄一共多少岁?

因此,由于未知数位置的变动,案例 7b 虽与案例 7a 结构相同,但问题元素互动程度降低,从而导致问题内在认知负荷(intrinsic cognitive load)下降。类似案例 7b 和 7a 的问题分别称之为关联性问题(connected problem)和非关联性问题(disconnected problem)(Tabachneck, Koedinger, & Nathan, 1995)。关联性问题中,两个变量之间的数量关系比较容易建立,较易于用算术方法求解。而在非关联性问题中,已知取值的变量之间没有直接的联结,无法直接建立关系。显然,非关联性问题在形成问题表征时需要整合的元素数量更多,因而难度更大。

与未知变量位置相关联的还有解决问题所需的运算类型。未知变量位置不同,会导致问题解决的运算类型的变化。仍以案例 7a 和 7b 为例,根据图 2 给出的代数式和已知条件,解决案例 7a 需要用到 $+$ 、 $\times$ 、 $/$,或者 $+$ 、 $/$,而解决案例 7b 需要用到 $+$ 和 $\times$ 。不同运算会导致问题难度的不同。研究表明,同等情况下, $+$ 和 $\times$ 运算相对容易,而 $-$ 和 $/$ 运算则难度较大(Nesher & HersHKovitz, 1994)。

需要推断的变量关系或信息 影响应用题解决的另一个因素是变量及其关系的明确程度。图 1 表明,案例 1 的 5 个三阶组合中只有两个是可以从问题陈述中直接转译的,其余 3 个则需要个体根据问题情境、世界知识和数学知识加以推断。例如,案例 1 并没有明确阐述在追及过程中,丁丁所走距离是丁丁行走速度和所用时间的乘积。该三阶组合依赖于个体对问题情境的理解,以及对“距离=速度 $\times$ 时间”这一数学公式的提取。如果个体缺乏相应的数学知识,则该三阶组合无法形成,问题也就无法解决。类似的,个体需要根据问题情境和已有经验推断爸爸追上丁丁时,两人所走的距离是相等的。与可以直接转译的变量关系相比,需要推断的变量关系需要调用更多的认知资源。问题初始状态中需要推断的信息越多,任务也就越复杂。因此,需要推断的三阶组合(或分步等式)数量(number of partial equation to be inferred)也是影响问题难度的一个因素(Arendasy

& Sommer, 2007)。

与变量关系推断数量相关联的是个体解题过程中涉及的数学概念、原理和公式的来源、数量和水平。Embretson 和 Daniel (2008)区分了数学问题解决中运用数学知识和公式的几种情况:(1)从问题陈述中转译(translate)数学公式,(2)回忆(recall)数学知识或公式,(3)为当前问题生成(generate)数学公式。对数学知识和公式的回忆,她们还区分了回忆数量以及知识所在年级水平。研究表明,数学知识和公式的来源(即转译、回忆和生成)显著影响问题解决的难度(Embretson & Daniel, 2008),需要个体回忆或者生成数学公式的问题难度显著增加(Daniel & Embretson, 2010)。数学知识所在年级水平以及回忆数量效应并不显著(Embretson & Daniel, 2008)。但是她们的研究采用了大学生作为被试,中学阶段的数学知识或公式的效应不显著可能是天花板效应。因此,该结论有待进一步验证。

问题结构与图式知识 由于数学问题的难度是任务特征与认知特征交互作用的结果,一个重要的因素是考察个体是否具备了特定问题类型对应的图式知识。问题图式的存在和激活使得原本复杂的问题表征和解决过程变成了模式识别和提取的过程,消解了上述各种因素所导致的问题复杂性,降低了问题内在认知负荷。因此,在分析代数应用题任务特征蕴含性的时候,必须同时考虑问题图式的存在和自动化程度(Paas, Renkl, & Sweller, 2003)。

尽管存在众多检测个体图式知识的手段和技术(Low & Over, 1992),但系统全面考查个体图式的类型、分化程度以及自动化水平是一项非常困难和复杂的任务。因此,有必要寻找一种简便易行的指标来指代问题图式的存在和自动化水平。专家-新手研究表明,图式知识的形成在很大程度上依赖于个体已有经验和教育经历(Chi et al., 1981; Feldon, 2007)。从这个角度考虑,一种可能的指标是考察个体已有教育经历中各种问题类型出现的频率。基于图式形成的曝光率假设(exposure hypothesis; Mayer & Tajika, 1991),个体在已有经历中接触某种问题类型的机会越多,相应问题图式形成的可能性越大,自动化水平也应该越高。已有研究证据表明了这种解释的可行性。Mayer (1981)分析了美国中学课本中代数应用题的类型

和发生频率,并在后继研究中(Mayer, 1982)发现问题类型的发生频率与个体对问题的回忆(recall)和重构(reconstruction)成绩之间存在很高的相关($r=0.85$)。这表明,问题所属类型的发生频率可作为指代问题图式存在的一种可能指标。在实际应用该指标时,如何确定某个具体问题所属的问题类型需要明确界定。在Mayer (1981)提出的代数应用题分类框架中,一个问题可以在族(family)、类(category)和模板(template)三个不同水平上进行类型划分。同一族内的不同问题享有共同的基本公式(source formula),如“距离=速度 $\times$ 时间”等。同一族的问题又可以根据表面故事的一般形式(general form of the story line)进一步划分为不同的类,如“行程”、“工作”、“混合”等。每类问题又可以根据问题背后具体的命题结构(proposition structure)划分为不同的问题模板,如“行程”中的“追及问题”。案例1就属于“行程追及”模板的一个变式。研究表明,在模板水平上划分问题类型能够较好的预测代数应用题的难度²(Arendasy & Sommer, 2007; Blessing & Ross, 1996)。具体应用时,可以直接利用问题所属模板的发生频率作为问题图式存在的指标。对项目生成研究而言,问题模板可以作为生成代数应用题的基础之一。而在运用离散型认知诊断模型进行诊断时,可以用问题模板及其发生频率作为指代问题图式的属性特征。通过个体对不同复杂程度和发生频率的问题模板的解答情况推断个体对应图式知识的存在和自动化水平。³

² 这些研究者根据同一模板的问题往往具有相近的表面故事这一现象,利用问题表面故事的典型性作为实际指标。问题表面故事的典型性是根据某一问题模板中某种表面故事的相对频率将其划分为典型(typical)、一般(neutral)和不典型(untypical)。但是确定三种情况的标准缺乏较为客观的依据。

³ 这里需要提醒的是,问题模板类型及其发生频率是受现有教材中代数应用题分布影响的,而作为项目生成的代数应用题则是超越现有教材的,是就整个代数应用题领域而言的,两者不可等同。此外,必须要意识到问题模板只是代数应用题的特定结构,而个体问题图式知识还包括了对应于该结构的数学知识、解题策略、世界性知识等等。因此,在没有个体问题解决认知过程信息的情况下,从问题模板的解答情况推断个体图式知识时必须谨慎下结论。

4.3 问题情境

问题情境是指代数应用题涉及到的对象、事件和场景,通常称之为问题的表面故事(surface story; Mayer, 1981)、表面结构(surface structure; Chi et al., 1981)或表面内容(surface content; Blessing, & Ross, 1996)。然而,从实质意义上看,问题情境并非可有可无的问题特征,而是提供了具体的问题背景(problem context),赋予了问题蕴含的各种变量及其结构的现实基础和意义(Koedinger, Alibali, & Nathan, 2008)。问题情境的变动情况既深刻影响着个体问题解决的认知过程,也对个体数学知识、技能和思维能力的形成至关重要。

Koedinger 和他的同事(Koedinger & Nathan, 2004; Koedinger, Alibali, & Nathan, 2008)研究了不同外部表征方式(external representation)对代数应用题解决认知过程的影响。研究发现,当向被试呈现相同数学结构的代数式(algebraic equation)、文字题(word problem)和故事题(story problem)时,个体在后两者的正确率高于前者。Koedinger 等人将这种现象解释为后两者以语言形式呈现而带来的熟悉效应(familiarity effects)。他们认为,代数应用题使用了个体熟悉的词汇,指向的是个体熟悉的对象或事件,因而提供了比符号代数式更多的经验基础(more grounded)。相比之下,代数式虽然更为简洁,但同时也舍弃了对熟悉对象和事件的参照。除非受过正式的代数训练,否则这些符号与应用题的故事陈述相比,个体会感到更加陌生。正是由于这种熟悉效应,从而使得个体能够根据有关问题情境的具体知识而形成一些非正式策略来解答问题,而不是通过抽象的代数符号式解答。

此外,应用题提供了更加丰富具体的语义描述(redundant semantic elaboration)和语义线索(semantic cues)来支持和验证个体的表征和推理,使个体更容易发现运算中的错误。例如,如果案例1中丁丁行走速度为100公里/分钟,个体容易觉察出步行是不可能这么快的。相比之下,单纯给出抽象符号式(abstract symbolic representation) $S=100$ 公里/分钟,则无法检验数据的现实性。

需要指出的是, Koedinger 等人所指的熟悉效应并不是一种情境效应(situational effects),而是言语陈述方式带来的语言优势(verbal advantage)。如

果是情境效应的话,那么故事题应该比单纯的文字题更加容易,然而他们发现两者在难度上没有显著区别,但又都比代数式简单。这一现象得到了其他研究的支持。Enright 等人(2002)向个体呈现结构相同的两种概率题。一种是嵌套在现实情境中的,另一种是单纯的文字题。结果发现两者在难度上没有显著区别。然而,如果同样的问题分别陈述为概率题或百分数问题的话,则发现概率问题比百分数问题更加困难。Enright 等人认为这种现象是因为百分数问题更为常见而导致的差异。

根据熟悉效应不是情境效应的假设,或许会得出量化结构相同但情境不同的代数应用题难度相同的推论。事实上这个推论是不成立的。由于个体对不同的问题情境的熟悉程度是不同的,因而,情境的不同也会导致问题难度的变化。Sebrechts 等人(1996)发现,涉及到金钱(money)的问题相对其它内容的问题更简单,并将其归因为个体具有与该类问题情境相关的大量经验。

除了熟悉效应之外,故事情境不同而导致的问题陈述中所涉及的概念(concept)、特征(attribute)以及关系(relation)的不同也会影响个体对问题相似程度的判断以及解题过程的迁移程度。通常,深层结构相同但情境不同的问题称为同形异态问题(isomorphic problems; Eysenck & Keane, 2000)。Reed (1987)研究了被试在两个同形异态代数应用题之间解题方案的迁移程度。研究发现,当两个同形异态问题都属于工作问题(work problem)时,迁移正确率为 76%,但当两个同形异态问题都属于混合问题(mixture problem)时,相应的正确率只有 38%。Reed 认为,这是由于在不同问题类型下,两个同形异态问题的表面故事中涉及到的概念在相似程度上有所不同,从而导致属于混合问题的两个同形异态问题的匹配透明度(mapping transparency)下降。匹配透明度是指个体觉察到一个问题中的概念和关系与另一个问题中的概念和关系相对应的容易程度。在工作问题中,两个同形异态问题都属于速度-时间问题,迁移时只需将两个问题的速度与速度、时间与时间分别相匹配。而在混合问题中,两个同形异态问题中一个涉及到溶液浓度,另一个涉及到利率。迁移时除了需要将溶质在溶液中所占的百分比与利息在本金中

所占百分比进行匹配之外,还需要将溶液重量与金钱总量相匹配。显然,两个工作问题的概念匹配相对容易,从而使问题解决方案在两者之间的迁移易于实现。此外,该研究还发现,同样情况下,个体更容易发现两个工作问题属于同形异态,从而从另一个角度证实了这种解释的可信性。

在有些情况下,问题情境对问题解决的影响作用非常微妙,取决于问题情境具体的呈现方式。例如,案例 8a 和 8b 除了问题陈述不同之外,在数学结构和问题情境方面完全相同(Hudson, 1983)。

案例 8a 有 5 只鸟和 3 条虫。假设这些鸟都来抢虫吃,每个鸟都想吃到一条虫。有几个鸟会吃不到?

案例 8b 有 5 只鸟和 3 条虫,鸟比虫多几个?

然而,研究发现,绝大多数幼儿园大班儿童和一年级小学生能够答对案例 8a (幼儿园大班: 83%; 一年级学生: 100%),但是只有约五分之一的大班儿童和约半数一年级学生答对案例 8b (幼儿园大班: 17%; 一年级学生: 64%)。Hudson 认为,这种差异是由于儿童在理解“……比……多几个”时的语言困难导致的。然而, Hudson 同时也提到另一种假设,即案例 8a 的陈述方式有可能更加符合儿童的生活经验。这种陈述方式本身提供了解决问题的线索,向个体建议了一种一对一的匹配或配对策略(one-to-one matching or pairing strategy)。这种解释也得到了其他研究的支持(Carraher, Carraher, & Schieman, 1987; Koedinger & Nathan, 2004; Koedinger, Alibali, & Nathan, 2008)。它表明在某些情况下,问题情境的呈现方式本身存在某些启动特征(enabling feature),能够深刻的影响个体对问题结构的理解和对问题解决策略的选择,进而影响问题解决结果。

根据问题情境与量化结构在具体问题中的不同关系,问题情境的影响效应不仅仅是促进性(facilitating)的,也有可能是抑制性(inhibiting)的(Hall, Kibler, Wenger, & Truxaw, 1989)。例如,案例 9a 和 9b 在深层结构相同,在问题情境方面也是相近的,都属于“工作”类型中的“合作”模板。按照 Reed (1987)的结构匹配模型,两个问题的基本公式框架都可以写成“ $R1 \times T + R2 \times T = \text{工作总量}$ ”的形式。对于案例 9a 而言,这一问题图式的激活有助于个体正确的写出问题解答算式“ $320 \times T +$

430×T=2400”。然而对案例 9b 而言,如果机械套用这一公式,则导致错误解答。个体需要在适当的抽象水平上将图式知识正确转化成适合于当前问题的信息。例如,相关图式中甲乙工作量相加在两个问题中都是适用的。但在案例 9b 中,个体需要将甲或乙的工作时间转化成他们相应的工作速率,然后套用问题图式的基本公式。这里,确定速率的合理表达方式是重要的一环,需要理解“工作总量”为抽象的 1 的问题情境使得相应图式知识在案例 9b 中应用的难度增加。

案例 9a 甲乙两个工程队合修一条长 2400 米的公路。甲队每天修 320 米,乙队每天修 430 米。两队从两端同时开工,几天后可以修完?

案例 9b 一件工作,甲单独做 20 小时完成,乙单独做 12 小时完成。如果甲、乙两人合做,需几小时完成?

5 结语:交互观下的代数应用题解决与后继项目生成研究的问题

本文在理论驱动的测量学范式下对代数应用题的认知加工过程及其任务特征进行了系统分析。该范式强调以测验项目的认知任务分析为桥梁,对测量建构的认知研究和测量学技术加以整合。就代数应用题而言,它强调问题解决过程的两个不同但彼此依存的方面。一方面是个体在代数应用题解决过程中表现出来的信息结构和认知特征,另一方面则是使得个体认知活动得以表现的任务特征。代数问题解决因而符合 Greeno (1994) 所指的代理者-情境交互作用过程(processes of agent-situation interaction)。在这一过程中,个体解决问题的实际表现既不是由个体认知特征决定的,也不是由任务特征决定的,而是两者共同决定的。如果把个体在问题解决过程中体现出来的认知特征视为代数问题解决能力的话,那么后者就是 Greeno (1994)所指的情境蕴含性(affordance of the situation)。每个代数应用题在本质上都创设了一个具体情境,而问题中所包含的特定任务特征则形成了该问题特定的限制条件(conditions for constraints),提供了个体特定信息结构或认知特征得以表现的前提和可能性。这种可能性未必一定变成现实。它转化成现实性的前提是个体具备适应于问题所提供的限制条件(attunement to constraints)的特定认知特征。由于不同问题所包

含的任务特征不同,因而在使不同代数问题解决能力得以表现的蕴含性也有所不同。

交互观下的代数应用题解决指明了项目生成研究可能存在的误区以及后继研究所需解决的问题。首先,它明确表明代数应用题中的任务特征与个体解决代数应用题的认知特征有着本质上的区别。认知特征是项目生成所试图指向的测量建构的内在表征,而任务特征只是代数应用题本身的特征。因此,如果仅仅在任务特征水平上对代数应用题进行分析,系统变动任务特征生成不同的代数应用题测验项目,虽然可以在任务特征水平上解释项目难度变化的来源,但却违背了认知项目设计的初衷,丧失了使测验项目设计成为建构理论驱动下的项目生成的可能性。其次,认知项目生成的关键环节是建立和标定项目任务特征与项目反应机制中认知特征之间的关系。本文在代数应用题认知加工过程的基础上分析了不同认知环节可能对应的任务特征。这种分析仅仅为后继代数应用题项目生成研究提供了可能的研究假设。后继研究需要通过实证手段,研究探索不同任务特征在代数应用题中的变化及组合方式对项目反应机制的影响,并通过测量建模的手段对这种影响力度进行标定。比较理想的方式是通过因子实验设计的手段,使得不同任务特征可以在各种代数应用题中进行独立的正交变换。然而在现实中,不同任务特征之间往往存在一定的关联性,或者同一任务特征可能对一个或多个认知过程产生影响。因此,实际研究中往往需要借助于统计建模手段来剖析不同任务特征对不同认知过程或成分的影响及力度。第三,与传统认知心理学研究不同,指向项目生成的实证研究不仅关注任务特征与认知成分之间关系的一般规律,同时关注这种关系在任务领域中的可推广性。如果代数应用题的认知加工模型只能充分解释某些问题类型,或者基于任务特征而生成的项目仅仅代表了代数应用题领域的某个子集,后继研究需要进一步探索该子集与其他代数应用题在任务特征及项目解决认知机制上的区别。这种分析不仅对于后继题库的建设是必要的,而且对于代数应用题认知模型的修改和完善也具有重要意义。它恰恰表明理论驱动的代数应用题项目生成研究是代数应用题解决能力的理论构建和推广不可或缺的手段。

参考文献

- Abedi, J. (2004). The no child left behind act and English language learners: Assessment and accountability issues. *Educational Research*, 33, 4–14.
- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, 14, 219–234.
- Alves, C. B., Gierl, M. J., & Lai, H. (2010). *Using automatic item generation to promote principled test design and development*. Paper presented at Paper presented at the annual meeting of the American Educational Research Association, Denver, CO, USA.
- Anderson, J. R. (1993). Problem solving and learning. *American Psychologist*, 48, 35–44.
- Arendasy, M., & Sommer, M. (2007). Using psychometric technology in educational assessment: The case of a schema-based isomorphic approach to the automatic generation of quantitative reasoning items. *Learning and Individual Differences*, 17, 366–383.
- Blessing, S. B., & Ross, B. H. (1996). Content effects in problem categorization and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 792–810.
- Bouwmeester, S., & Sijtsma, K. (2004). Measuring the ability of transitive reasoning, using product and strategy information. *Psychometrika*, 69, 123–146.
- Carraher, T. N., Carraher, D. W., & Schlieman, A. D. (1987). Written and oral mathematics. *Journal for Research in Mathematics Education*, 18, 83–97.
- Carroll, J. B. (1993). *Human cognitive abilities – A survey of factor analytic studies*. Cambridge: Cambridge University Press.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.
- Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review*, 76, 387–404.
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American psychologist*, 12, 671–684.
- Daniel, R. C., & Embretson, S. E. (2010). Designing cognitive complexity in mathematical problem-solving items. *Applied Psychological Measurement*, 34, 348–364.
- De Corte, E., Verschaffel, L., & De Winn, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology*, 77, 460–470.
- Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93, 179–197.
- Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods*, 3, 380–396.
- Embretson, S. E. (2004). The second century of ability testing: some predictions and speculations. *Measurement*, 2, 1–32.
- Embretson, S. E., & Daniel, R. C. (2008). Understanding and quantifying cognitive complexity level in mathematical problem solving items. *Psychology Science Quarterly*, 50, 328–344.
- Embretson, S. E., & Gorin, J. (2001). Improving construct validity with cognitive psychology principles. *Journal of Educational Measurement*, 38, 343–368.
- Embretson, S. E., & Yang, X. (2007). Automatic item generation and cognitive psychology. In S. Sinharay & R. Rao (Eds.), *Handbook of Statistics: psychometrics* (pp. 747–768). Amsterdam, The Netherlands: Elsevier, B.V.
- Enright, M. K., Morley, M., & Sheehan, K. M. (2002). Item by design: The impact of systematic feature variation on item statistical characteristics. *Applied Measurement in Education*, 15, 49–74.
- Eysenck, M., & Keane, M. T. (2000). *Cognitive Psychology: A student's handbook*. (4th ed.). Philadelphia: Psychology Press.
- Feldon, D. F. (2007). The implication of research on expertise for curriculum and pedagogy. *Educational Psychology Review*, 19, 91–110.
- Gierl, M. J., & Haladyna, T. M. (2012). *Automatic item generation: Theory and practice*. New York, NY: Routledge.
- Gorin, J. S. (2005). Manipulating processing difficulty of reading comprehension questions: The feasibility of verbal item generation. *Journal of Educational Measurement*, 42, 351–373.
- Greeno, J. G. (1994). Gibson's affordances. *Psychological Review*, 101, 336–342.
- Hall, R., Kibler, D., Wenger, E., & Truxaw, C. (1989). Exploring the episode structure of algebra story problem solving. *Cognition and Instruction*, 6, 223–283.
- Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, 87, 18–32.
- Hinsley, D., Hayes, J. R., & Simon, H. A. (1977). From words to equations-Meaning and representation in algebra word problems. In P. Carpenter & M. Just (Eds.), *Cognitive Processes in Comprehension* (pp. 89–106). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Horn, J. L., & Cattell, R. B. (1966). Refinement and test of theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, 57, 253–270.

- Hudson, T. (1983). Correspondence and numerical differences between disjoint sets. *Child Development*, 54, 84–90.
- Irvine, S., & Kyllonen, P. (2002). *Generating items for cognitive tests: Theory and Practice*. Mahwah, NJ: Erlbaum.
- Johnson, E., & Monroe, B. (2004). Simplified language as an accommodation on math tests. *Assessment for Effective Intervention*, 29, 35–45.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, 92, 109–129.
- Kintsch, W., & Lewis, A. B. (1993). The time course of hypothesis formation in solving arithmetic word problems. In M. Denis & G. Sabh (Eds.), *Modeles et concepts pour la science cognitive: Hommage a Jean-Francois Le Ny* (pp. 11–23). Grenoble, France: Press Universitaires.
- Koedinger, K. R., Abibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science*, 32, 366–397.
- Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences*, 13, 129–164.
- Lane, S. (1991). Use of restricted item response models for examining item difficulty ordering and slope uniformity. *Journal of Educational Measurement*, 28, 295–309.
- Lewis, A. B. (1989). Training students to represent arithmetic word problems. *Journal of Educational Psychology*, 81, 521–531.
- Lewis, A. B., & Mayer, R. E. (1987). Students' misconception of relational statements in arithmetic word problems. *Journal of Educational Psychology*, 79, 363–371.
- Loftus, E. F., & Suppes, P. (1972). Structural variables that determine problem-solving difficulty in computer-assisted instruction. *Journal of Educational Psychology*, 63, 531–542.
- Lohman, D. F. (2000). Complex information processing and intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 285–340). New York: Cambridge University Press.
- Lohman, D. F., & Ippel, M. J. (1993). Cognitive diagnosis: From statistically based assessment toward theory based assessment. In N. Frederiksen, R. J. Mislevy, & I. I. Bejar (Eds.), *Test theory for a new generation of tests* (pp. 41–71). Hillsdale, NJ: Erlbaum.
- Low, R., & Over, R. (1992). Hierarchical ordering of schematic knowledge relating to area-of-rectangle problems. *Journal of Educational Psychology*, 84, 62–69.
- Marshall, S. P. (1995). *Schemas in problem solving*. Cambridge: Cambridge University Press.
- Mayer, R. E. (1981). Frequency norms and structural analysis of algebra story problems into families, categories and templates. *Instructional Science*, 10, 135–175.
- Mayer, R. E. (1982). Memory for algebra story problems. *Journal of Educational Psychology*, 74, 199–216.
- Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science*, 26, 49–63.
- Mayer, R. E., Larkin, J. H., & Kadane, J. B. (1984). A cognitive analysis of mathematical problem-solving ability. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 231–273). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mayer, R. E., Tajika, H., & Stanley, C. (1991). Mathematical problem solving in Japan and the United States: A controlled comparison. *Journal of Educational Psychology*, 83, 69–72.
- Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra-word-problem comprehension and its implications for the design of learning environments. *Cognition and Instruction*, 9, 329–389.
- Nesher, P., & HersHKovitz, S. (1994). The role of schemes in two-step problems: Analysis and research findings. *Educational Studies in Mathematics*, 26, 1–23.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Paas F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38, 1–4.
- Ree, M. J., & Carretta, T. R. (1995). Group differences in aptitude factor structure on the ASVAB. *Educational and Psychological Measurement*, 55, 268–277.
- Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 13, 124–139.
- Roid, G. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside Publishing.
- Sebrechts, M. M., Enright, M., Bennett, E., & Martin, K. (1996). Using algebra word problems to assess quantitative ability: Attributes, strategies and errors. *Cognition and Instruction*, 14, 285–343.
- Shafel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English

- language learners and students with disabilities. *Educational Assessment*, 11, 105–126.
- Shalin, V., & Bee, N. V. (1985). *Analysis of the semantic structure of a domain of word problems* (Tech. Rep. No. APS-20). Pittsburgh: University of Pittsburgh, Learning Research and Development Center.
- Snow, R. E., & Lohman, D. F. (1989). Implications of cognitive psychology for educational measurement. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 263–331). New York, NY: American Council on Education/Macmillan.
- Stillman, G. (1996). Mathematical processing and cognitive demand in problem solving. *Mathematics Education Research Journal*, 8, 174–197.
- Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 81, 457–466.
- Sweller, J., van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10, 251–296.
- Tabachneck, J. J. M., Koedinger, K. R., & Nathan, M. J. (1995). A cognitive analysis of the task demands of early algebra. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation.

Cognitive Processes and Task Features of Algebra Story Problems for Item Generation

YANG Xiangdong

(Institute of Curriculum and Instruction, East China Normal University, Shanghai 200062, China)

Abstract: Cognitive item generation, rooted from the theory-based measurement paradigm, requires formulating an elaborated theory of the measured construct, upon which task features that are linked to cognitive variables underlying such theory can be systematically analyzed and manipulated. Cast from the perspective of problem solving as an agent – situation interactive process, both cognitive processes of algebra story problem solving and task affordance of such problems are systematically analyzed in this article. Specifically, an algebra story problem is considered as a particular quantitative structure embedded in a verbally-stated realistic situation. Three different types of representations, i.e., a textbase, a situational model and a quantitative model, need to be formed during the four phases of algebra story problem solving including translation, integration, planning and execution, which are further mediated by the schematic knowledge possessed by the problem-solver. Correspondingly, Task features that are associated with syntactic/semantic relations, quantitative elements as well as their structures, and problem situations are analyzed and synthesized based on a broad range of relevant literature. It is evident from the current analysis that subsequent research should focus on empirical studies to link the set of cognitive variables and task features for such problems, as well as to evaluate the generalizability of such linkages across the targeted domain in order for item generation of algebra story problem to lay on a solidly scientific foundation.

Key words: theory-based measurement; item generation; algebra story problem; cognitive process; task feature