性ENRE 别与数学评价：评价事件管理的比较’EPISODES

BRISMONTIER[1*] Chloé

摘要 – 本文关注法国教育体系中义务教育的前四个月里便已显现的女生与男生在数学表现上的差异。文章探讨了性别刻板印象对教师评价实践的影响，以及性别刻板印象威胁对学生表现的影响。为寻找答案，本文借助不同的理论工具，对评价事件的管理进行比较研究。

关键词 ： 评价事件、性别刻板印象、调节、表现、情感

摘要 –本文关注法国教育体系中义务教育的前四个月里便已显现的女生与男生在数学表现上的差异。文章探讨了性别刻板印象对教师评价实践的影响，以及性别刻板印象威胁对学生表现的影响。为寻找答案，本文借助不同的理论工具，对评价事件的管理进行比较研究。

关键词：评价片段、性别刻板印象、调节、表现、情感

I. 引言

在法国，尽管全国性评估旨在提供精准的教育测量学和心理测量学方法，却未能解释女生与男生之间数学成绩差距的成因，而这一差距往往对女生不利。缺乏相关阐释，削弱了科学界反驳关于不同性别——即生理性别——之间存在智力差异的有争议假设的能力（C. Perronet et al., 2024）。教育科学、心理学和社会学等多个学科都提出了不同的解释。其中，教师的评价实践如今也受到审视。人们怀疑这些实践会受到性别刻板印象的影响，进而改变学生的数学情感；按照 V. A. Debellis 和 G. A. Goldin（2006）的理论，数学情感包括与数学相关的信念、价值观、情绪和表征，而这可能不利于女生的表现。这可能体现为根据学生性别而有所不同的评价中的教学契约。我们所说的评价中的教学契约，是指教师在知识学习过程中向学生提出的、针对某一评价环节的特定教学契约（Sayac, 2017）。差异化教学契约这一概念是我博士论文的核心研究对象，我将在论文中考察其特征，以及它与学生数学表现和数学情感之间的联系。本文首先对评价、预测与表现司（DEPP）全国性评估中的学生成绩进行图形分析，展示女生与男生之间的数学成绩差距。随后，我将回顾解读这一差距的主要科学研究，最后介绍我研究框架中的一部分，探讨性别对教师评价实践的影响，以及这些实践与学业表现、学生对数学的情感之间的关系。

[1*] 巴黎西岱大学 – 法国 – chloe.brismontier@etu.u-paris.fr

II. DEPP 全国性评估

自2018-2019学年以来，DEPP与教育总司（DGESCO）以及教育、体育与研究总督察局（IGESR）合作，设立了一项名为“基准评估”的全国性评估机制。该机制旨在评估法国学校学生在法语和数学方面的掌握水平，同时帮助教师调整课堂的教学组织。面向预备班（CP）、小学二年级（CE1）、CM1以及6ème 年级（6ème）学生的数学评估——分别对应英语学制中的一年级、二年级、四年级和六年级——是我研究的主要数据来源之一。

在我的数学教学法博士论文中，我特别关注小学（一年级至五年级）和六年级班级中女生与男生之间的成绩差距，即6至12岁学生之间的差距。我建议先按年级和数学任务类型，分析2018年至2023年「基准评估」的结果数据。然后，再按数学领域对结果进行图形分析。

*1. 2018*年至*2023*年按年级和数学任务类型划分的全国性评估结果

DEPP全国性评估机制的第一轮于2018年9月开展。最初，「基准评估」仅面向小学一年级、CE1和6e学生。随后，该机制于2023-2024学年扩展至CM1班级，并于2024-2025学年扩展至CE2和CM2班级。除小学一年级学生外，每名学生在学年初参加一次测评。小学一年级学生则需参加两次测评：一次在学年初，另一次在学年中。每轮「基准评估」结束后，DEPP都会发布信息简报，介绍对学生答卷进行统计处理后得出的结果。这些信息简报可在DEPP网站[2], 查阅，且附有一个数据文件，其中包含用于制作表格和图表的数据。借助DEPP发布的数据，我们可以了解部分全国性评估中，按年级和受评数学任务类型划分的女生与男生之间的成绩差距。

根据 DEPP 的小学一年级数据，在2018年9月、2019年、2020年和2023年CP年初进行的测评中，女生在几乎所有受评估的数学任务类型上都比男生表现更好。2018年至2022年，注册小学一年级的女生在数的比较和在数轴上标示一个数方面表现不如男生。相比之下，在2023年CP年初的测评中，女生在图形组合再现和问题解决方面表现不如男生。

2 https://www.education.gouv.fr/les-evaluations-nationales-exhaustives-307627

CP年中出现了一种特殊现象。自2020年以来，每次“基准评估”的结果都显示，在数学评估的所有任务类型上，CP年中的女生表现均不如男生。

对于就读小学二年级的学生，女生与男生之间的成绩差距（女生处于劣势）比小学一年级年初大得多，也更常见。CP年中观察到的现象在CE1年初再次得到证实。直到2022年，女生在图形组合再现和心算方面的表现优于男生（也仅限于这两项）。2023年，女生在加法和心算方面的表现优于男生。

根据DEPP提供的数据，对于2023年9月就读CM1的学生，女生中达到令人满意的掌握程度的比例仅在列式计算和数字计算方面高于男生。与数学其他受评任务类型中男生占优的成绩差距相比，这一有利于女生的成绩差距微不足道。

对于6e年初的全国性评估，DEPP未公布2018年的数据；此后各年的现有数据也没有按学生性别区分表现，2023年9月的测评除外。

*2. 2018*年至*2023*年各学段按数学领域划分的全国性评估结果

为了对不同学生群体在小学一年级、小学二年级、CM1和六年级全国性评估中的结果进行历史分析，我将通过图表解读，展示女生和男生在数学各评估领域达到令人满意的掌握程度的比例变化。

我说明一下，归入“数与计算”的数学任务类型，除六年级以这一名称评估的任务外，还包括以下内容：

- 比较整数（小学一年级），- 读整数（小学一年级和小学二年级），- 写整数；在数轴上标出一个数（小学一年级、小学二年级和CM1），- 进行加法；进行心算；进行减法；表示整数（小学二年级），- 列式并计算（CM1）。

对于“空间与几何”，以此名称在六年级进行评估的数学任务类型，以及其他年级的数学任务类型包括：

- 对集合计数（小学一年级），- 复制组合图形（小学一年级和小学二年级）。

关于“问题解决”，这里只考虑在 CP年初、小学二年级、CM1 和六年级进行评价的问题解决。关于“自动化技能”，目前只有六年级评价的数学任务类型，以及2023年 CM1 的数值事实和程序记忆数据。由于数据不足，无法绘制女生和男生在“自动化技能”领域达到令人满意的掌握程度的比例变化历史图。

因此，我的分析不包括这一领域。最后，CM1 的全国性评估不涉及“空间与几何”领域。

“数与计算”（图表1）和“问题解决”（图表3）领域的图表曲线呈现出相似性。在每个有数据的年份中，都可以观察到，CP年初女生和男生的学生表现优于小学二年级。2023年，从小学一年级到六年级，学生表现呈下降趋势。“数与计算”（图表1）和“问题解决”（图表3）中，女生达到令人满意的掌握程度的平均值曲线均低于男生，但走势相似。在“数与计算”中，女生学生曲线与男生学生曲线之间的平均差距为-3.3%；而在“问题解决”中，平均差距为-3.03%。

“空间与几何”领域的图表曲线（图表2）呈现出不同于前两个领域的变化轨迹。直到2022年，在“空间与几何”领域达到令人满意的掌握程度的男生和女生比例，随着年份变化以及从小学一年级到小学二年级的年级变化都很小。此外，达到令人满意的掌握程度的女生比例通常高于男生，但2023年就读小学二年级的女生除外。2023年测评结果显示，从小学一年级到小学二年级，以及从小学二年级到六年级，表现均大幅下降。

3. 结论

无论学生的性别以及针对学生所评价的数学领域如何，对于达到令人满意的掌握程度的学生而言，从小学一年级到六年级，学生在数学方面的表现均呈下降趋势。然而，值得注意的是，CP年初女生在数学方面的表现优于男生，这与小学二年级、CM1和六年级的结果相反，“空间与几何”领域除外。从教学法角度来看，这一观察引发了人们对全国性评估有效性的疑问（Grapin和Mounier，2020）。事实上，DEPP在其文档中明确指出，小学一年级、小学二年级和CM1的评估并不测量学生的数学能力，其目的也不是测量前几年所学的一切，而是测量学生的基本认知能力；这不同于面向六年级学生的评估。鉴于这一说明，我们可以质疑“基准评估”结果所描述的小学学生数学掌握水平。不过，这些结果揭示出，在数学评估的众多任务类型中，被DEPP评定为达到令人满意的掌握程度的女生比例与男生比例之间存在不可忽视的差异。

为解释女生和男生在“基准评估”中的表现差异，不同学科领域对此进行了多方面的探讨。我将对有关这一差距特征的信息简报、文章、著作和调查进行非穷尽性的综述。

III. 法国女生—男生表现差异特征的非穷尽性文献综述

近期，公共政策研究所（IPP）的经济学家发布了一份题为《女生在数学方面的掉队》（2024）的报告。然而，正如

DEPP提供的数据所显示的那样，数学表现下降的不仅是女生，男生也受到影响。不过，IPP的这项研究有助于更清楚地描述女生与男生之间不利于女生的数学成绩差距。首先，这一差距的变化主要影响CP年初表现最好的女生，并且在所有社会类别和家庭结构中，以及全国各地均有体现。其次，无论初始表现如何，女生的进步都不如起点水平相同的男生。最后，在女生占多数的课堂中、由女性而非男性担任教师时，以及学校位于优先教育网络升级版（REP+）区域时，这一差距较小。

通过 ELFE调查（法国儿童纵向研究），法国国家人口研究所（Ined）早在2022年就指出了女生与男生之间数学成绩差距的若干特征。Fischer 和 Thierry 撰写的新闻稿（2022年）指出，这种不利于女生的差距可能从5岁或6岁起就能观察到，且不受年级影响。作者认为，性别影响并非源于先天因素，但其成因仍无法解释。其他研究者也探讨了性别影响法国教育体系中学生表现这一假设。多位女性研究者，尤其是教育科学、社会学和数学教学法领域的研究者，对教师的评价实践提出质疑（Mathieu-Benmerah，2023年；Perronnet 等，2024；Sayac，2024），认为这些实践会再生产性别刻板印象。教师给予的反馈和评价可能因学生性别而异，并对女生不利。因此，女生的能力感和自我效能感可能会受到削弱（Imberdis 等，2021）。除了自我认知和自尊（Martinot，2001；Bishop，2014；Darragh 和 Radovic，2018）之外，学生的数学表现和动机信念也可能对其数学学业成功发挥作用。“学生会内化与各学科相关的某些群体图式”（Jamain，2019，第151页），这些图式与他们的能力感以及对各学科的兴趣密切相关（Imberdis 等，2021）。N. Mella（2021）进一步指出，学业成功与社会情绪能力及学校适应措施有关。她提出了一个尚未得到验证的假设：学校自我调节与社会情绪能力之间的关系，取决于学生运用有效学习策略的能力。此外，她指出，学生对评价任务重要性的感知因学科而异，而在数学中，这种感知可能不利于女生的学业成功（Mella，2021）。

有必要开展一项定性研究，探讨女生与男生之间不利于女生的数学成绩差距，以更好地理解学业成功、学生的数学情感与个体调节之间的依存关系。为考察这些关系，我选择主要以评价中的教学契约为研究框架（Sayac，2017）。这一框架提出了范式转变，立足课堂现实，同时考虑情境因素与个人因素，旨在描述并解释教学、学习及其评价之间互动的相关现象。我提出假设：学生性别不同，评价中的教学契约也存在差异。

IV. 我的研究理论工具的界定

为分析性别对评价中的教学契约的影响，本研究对“问题解决”和“空间与几何”学习课中观察到的非正式评价事件的处理方式进行了比较研究。选择这两个数学领域，是基于女生与男生之间表现的差异，图表2和3对此有所体现。为收集足够数量的数据，本研究在玛丽亚姆·米尔扎哈尼项目实验区域内的七所小学和三所初中开展，该区域隶属里尔学区教育局[3]。研究旨在考察互动调节以及评价中的专业与教学判断（JPDE）如何融入学生的学习策略，并如何随多个因素而变化：所布置的数学任务及其复杂度、教师的评价实践，以及学生的性别。为此，研究拟从四个层次进行分析：从评价中的教学契约视角分析评价事件；依据活动理论分析学生的数学活动；依据 L. Allal（2007）的概念框架分析学生的调节策略；以及分析教师与学生之间的互动调节。本文聚焦于第四个分析层次，详细介绍评价事件中观察到的互动调节类型，并将其与所布置的数学任务联系起来。

1. 互动调节的分类

所研究的非正式评价事件，是指在学习课中观察教师与学生之间日常互动的情形。我认为，这些互动以教师作出判断的能力或行为为基础，判断内容包括：

- 关于学生的数学知识、解题程序及结果，且 n 指向布置的数学任务或规定、教授的知识；- 具有指令性，并与评价事件的情境要素相关；- 关于学生的自我调节或学生本人。

第一种情况下，教师的判断提供的信息会影响学生数学活动的过程或维度，分别对应于 N. Grapin 和 É. Mounier（2024）所建模的问题解决活动，或 T. Barrier、C. Hache、A.-C. Mathé 和 S. Montigny（2013）所界定的几何活动。第二种情况下，教师的判断提供的信息涉及其评价中的专业与教学判断。最后，在第三种情况下，教师的判断提供的信息会影响学生的数学情感，以及其运用自我调节策略的能力（Debellis和Goldin，2006）。根据我对日常互动的描述，这些互动可能会因教师作出判断的能力或行为而引发互动调节。由此可以归纳出三类互动调节：教学类别、指令类别和情感类别。基于这一互动调节分类，我们将制定相应的分析标准，以研究其规律性与变异性，从而识别其中发挥作用的情境因素和

[3] https://filles-maths-nsi-projet-maryam-mirzhakhani.site.ac-lille.fr

个人因素，这些因素影响着问题解决和几何活动中评价环节的管理。

2. 互动调节的规律性与变异性分析

借鉴 M. Kiwan-Zacka 和 E. Roditi（2018）关于过程分析的研究（该研究为教师与一名或多名学生之间的每一次“教学调节”关联一组信息—行动），我将通过确定每个类别中可能出现的信息和行动，构建互动调节的分析标准。

对于教学类别，可能的信息—行动组合是根据 M. Kiwan-Zacka 和 E. Roditi（2018）对“教学调节”的定义界定的。

> “一次‘教学调节’的特征是：一名或多名学生在完成数学任务时产生某种表现，教师对其进行解读，并由教师采取干预措施；该干预可以是帮助完成任务，也可以是修改——简化——待完成的任务”（Kiwan-Zacka 和 Roditi，2018）。

为了区分在所观察的两个数学领域中，旨在帮助或尝试帮助完成数学任务的行为，与改变待完成任务的行为，我依据 E. Petitfour（2015）对工具化建构中所提供帮助的分类，并参考了基于描述教师如何帮助学生的文章（Croguennec，2023年；Veyrunes 等，2005）所构建的问题解决帮助分类。根据这项理论工作，可能的信息—行为组合涉及以下内容的组合：有关学生知识、程序或结果的信息，以及导致数学任务发生变化的行为，或帮助、尝试帮助求解数学任务的行为。

对于指令类别，互动调节可以与课堂情境相关的信息相联系，例如题目情境包装、关于课堂流程的指令以及时间安排，也可以与某种行为相联系：

- 阐明或修改评价标准和活动目标，- 修改题目情境包装，- 确认，- 追问，或重新表述问题或学生的回答，- 要求一名学生帮助另一名学生；- 比较两名学生的作品，或将一名学生的作品与参照标准进行比较。

上述行动是独立于所提出的数学任务的行动，其中所教授或规定的知识被视为参照标准，且在数学任务的求解过程中不对学生进行指导。

对于情感类别，互动调节可以与以下信息相关联：学生的局部情感特征——例如情绪、态度或动机；学生的整体情感特征——例如自尊、信念或表征；或者学生调节情绪的能力（出于好奇而行动，将挫折感视为调整策略的信号

，调动积极情绪以投入数学任务，减轻或消除消极情绪以专注于数学任务等）。教师可以根据对这些信息的解读，对学生的情感能力和情感结构采取行动（Debellis和Goldin，2006），也可以影响学生对数学的自尊、自我效能感和信念（Hattie和Timperley，2007）。

教师采取的每一项行动都可能以不同程度的意识和意向性进行，从而导致学生调节机制的改变或强化（Allal，2007）。在本比较研究中，必须考虑学生调节机制的状态及其因教师介入而产生的变化。事实上，通过比较学生的调节类型，可以观察哪些“互动模式”会根据数学任务促进或不促进学生能力的转变，以及这种转变是否因学生性别而有所差异。如果互动调节能够实现以下目标，则将被视为有效：

> “引导、控制或调整学习者的认知、情感和社会活动，从而促进学习者能力的转变”（Allal，2007）。

否则，这些策略将被视为无效。P. A. Genoud 和 M. Guillod（2014）通过一份评估数学学习社会情感态度的问卷（共有416名学生参与）表明，学生的信念与其情绪密切相关，而情绪会影响认知过程的启动。这一发现说明，关注“互动模式”对学生及其待掌握技能的影响具有重要意义，这可以作为性别刻板印象威胁影响学生数学学习的指标。

V. 结论

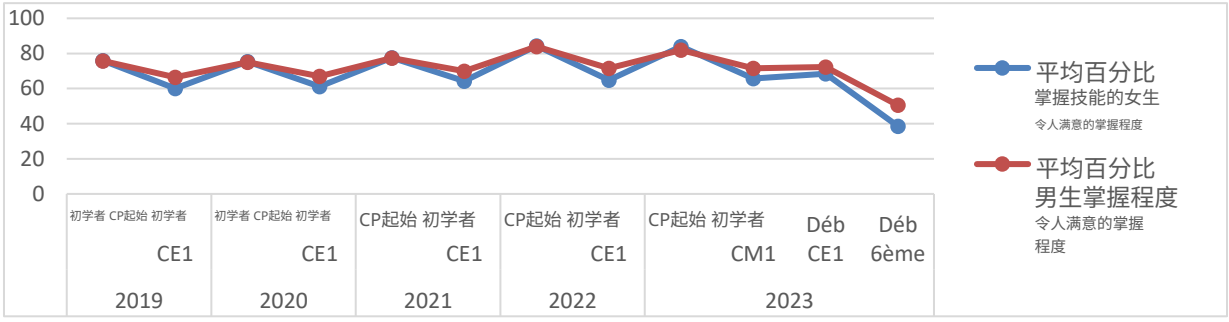
本文撰写于我开展数学教学法博士论文研究的初期，探讨了在数学活动中，如何把握学生表现层面实际发生的情况这一问题。许多定性研究关注性别刻板印象威胁对数学学习的影响、性别刻板印象对教师评价实践的影响，以及学生根据性别作出的教育方向选择。然而，针对某项具体能力或数学任务进行分析的研究寥寥无几。这一发现促使我结合数学教学法、认知心理学和数学情感领域的不同理论工具，从两个研究方向展开探讨：性别刻板印象对教师评价实践的影响，以及性别刻板印象威胁对学生表现的影响。我的研究项目旨在探究教师的评价实践、学生表现及其数学情感之间的依存关系，以评估评价中的教学契约是否会因学生性别而有所不同。

参考文献

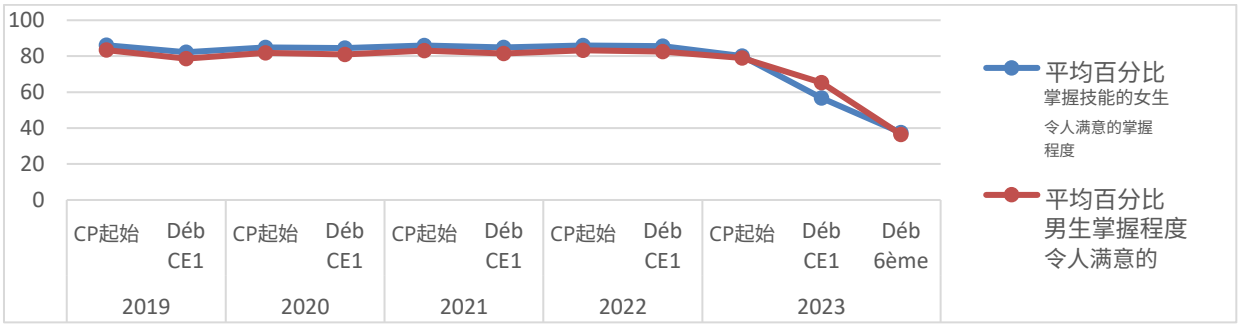
Allal, L. (2007). 引言。学习调节：教育研究与实践的概念方向。载于学校情境与培训中的学习调节（第 7-23 页）。De Boeck Supérieur；Cairn.info。Barrier, T., Hache, C., Mathé, A.-C., & Montigny, S. (2013). 描述学生的几何活动：工具、视角与语言。COPIRELEM，法国南特。Bishop, A. (2014). 数学教育中的价值观。载于 S. Lerman（编），数学教育百科全书（第 633-636 页）。Springer Netherlands。Croguennec, F. (2023). 小学问题情境的解决：教学差异化与教学法的挑战。《教育科学评论》，*49*(1)。Darragh, L., & Radovic, D. (2018). 数学学习者身份认同。载于 S. Lerman（编），数学教育百科全书（第 1-4 页）。Springer International Publishing。Debellis, V. A., & Goldin, G. A. (2006). 数学问题解决中的情感与元情感：一种表征视角。数学教育研究，*63*(2)，131-147。JSTOR。DEPP。（2024年6月）。全国性全面评估 [政府网站]。Education gouv。https://www.education.gouv.fr/les-evaluations-nationales-exhaustives-307627 Fischer, J.-P., & Thierry, X. (2022). 男生的数学表现与女生相比，在 6 岁时跃升（至少在 ELFE 数据中如此）。英国发展心理学杂志，*40*，504-519。Genoud, P. A., & Guillod, M. (2014). 设计并确认一份评估数学社会情感态度的问卷。教育研究，*20*。Grapin, N., & Mounier, E. (2020). 关于法国义务教育初期全国性评估的教学法观点。学校数学评论，*234*，22 30- 。Grapin, N., & Mounier, É. (2024). 评估小学一年级学生解决加法问题的能力：一项围绕数值大小展开的探索性研究。*Grand N*，*113*，29-48。Hattie, J., & Timperley, H. (2007). 反馈的力量。*77*(1)，81-112。Imberdis, A., Toczek, M.-C., Dutrévis, M., & Sacré, M. (2021). 学校学科与性别刻板印象：学生与教师的看法。学校与职业方向指导，*50*(4)。Jamain, L. (2019). 法语和数学能力的自我评估偏差：小学学生中的发展变化及其对学生社会心理适应和学业成功的影响？ [教育]。格勒诺布尔-阿尔卑斯大学。Kiwan-Zacka, M., & Roditi, É. (2018). 教学调节与学生学习。法语数学空间 *2018*，77-85。Martinot, D. (2001). 自我认知与自尊：学业成功的要素。《教育科学评论》，*27*(3)，483-502。Érudit。Mathieu-Benmerah, M. (2023). 评价片段与平等：小学数学与体育与运动教育中的教学法比较。性别、教育与培训（*GEF*），*7*。Mella, N., & al. (2021). 青春期的社会情感能力与学业表现：学校适应有何作用？心理学前沿，*12*(640661)。日内瓦大学开放档案库。Perronnet, C., Marc, C., & Paris-Romaskevich, O. (2024). 数学女学者：女生，数学的未来。法国国家科学研究中心出版社。

Petitfour, E. (2015年8月). 工具化建构中所提供帮助的分类. EEDM18：第18届数学教学法暑期学校，布雷斯特，法国。Sayac, N. (2017). 数学评价及其实施的教学法研究：学习与培养的关键问题 [教育]. 巴黎狄德罗大学—巴黎第七大学。Sayac, N. (2024). 如何解释小学一年级和小学二年级全国性评估中女生与男生数学表现的差异？《评价——国际教育与培养研究期刊》，*9*(3), 32-51. Veyrunes, P., Durny, A., Flavier, E., & Durand, M. (2005). 教师与学生的活动如何衔接，以便在小学解决数学问题：一项案例研究。《教育科学评论》，*31*(2), 471-489.

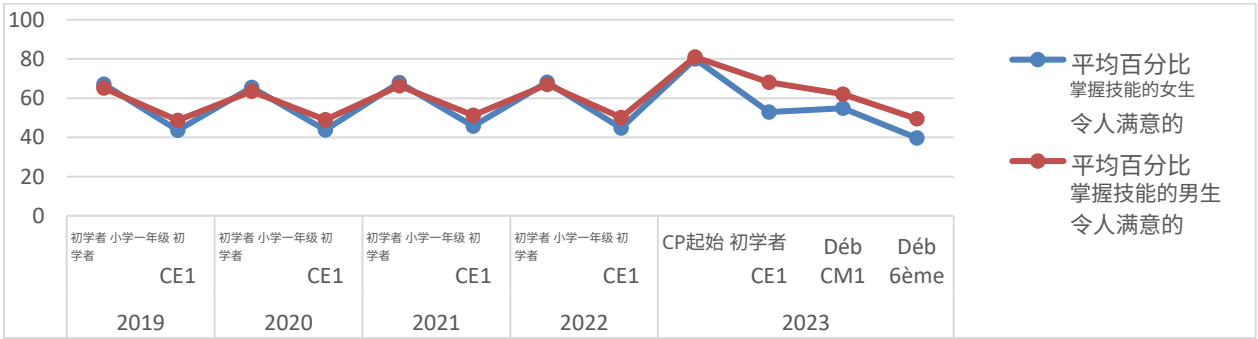
附录



图表 ***1*** *- 2018至2024年「数与计算」令人满意的掌握程度中女生与男生成绩差异*



图表 ***2*** *- 2018至2023年「空间与几何」令人满意的掌握程度中女生与男生成绩差异*



图表 ***3*** *- 2019至2024年「问题解决」令人满意的掌握程度中女生与男生成绩差异*