

## 《心理学报》审稿意见与作者回应

题目：3~8 岁儿童对群体认知的社会性线索偏好及发展特点

作者：王阳，温芳芳，佐斌

---

### 第一轮

#### 审稿人 1 意见：

论文采用质性和量化研究相结合的方式考察了 3-8 岁儿童群体分类线索偏好的发展特点，研究选题是发展心理学和社会心理学相关研究领域长期关注的重要问题。论文研究设计合理，数据统计分析得当，对于群体可视线索和本质线索的区分具有良好的理论依据和理论创新。论文总体写作规范，逻辑清晰。

**意见 1：**引言部分，作者重点介绍了可视线索与本质线索的区分及其理论依据，但对于研究的两个核心假设论证不足。儿童为什么会在社会群体分类中更加偏好群体本质线索？为什么随着年龄的增长，儿童对于群体本质线索的偏好会逐渐增强？其背后的认知发展机制是什么？

**回应：**非常感谢您的宝贵建议，我们针对您的问题进行了进一步思考与修改。在最新版本的论文中，我们补充了以往相关的研究结果及背后可能的认知机制，完善了研究假设。在这次写作修改中，基于已有前人的相关理论和研究发现，结合两位审稿人的意见，我们将群体“可视线索”与“本质线索”两种类型的命名修改为“物质性线索”和“社会性线索”。同时，采用“物质性线索”和“社会性线索”的概念与命名阐述和修改了全文的相关内容。

首先，尽管以往研究没有从物质性和社会性视角讨论群体线索，但依然有相关的理论和实证研究可以为这两类线索的提出提供一定的参考和支持。社会分类的直觉理论提出儿童存在两种理解社会类别的系统，一种将群体视为自然类别基础上产生的，这类群体线索往往和生理特征有关；另一种则是关注群体内部成员间的社会义务，认为群体内部更可能存在心理和行为上的相互支持。Rhodes 等人继而提出，相比群体在外部可视信息上的相同或差异性，群体内部成员间的社会互动模式在儿童的群体认知中扮演更重要的角色。实证研究层面上同样发现了儿童用群体归属预测成员关系的能力同样在 1 岁左右即有所体现，也有研究通过对比儿童对类别标签和类别规范特质的依赖性发现，类别规范特质在 4-5 岁儿童的社会认知中扮演更重要的地位。因此，在已有理论和实证研究基础上，本研究假设儿童在群体分类中更加偏好使用社会性线索而非物质性线索。

另一方面，关于儿童社会性线索偏好发展这一问题的假设则是来源于以往有关具有视觉显著性的线索（如性别、种族和年龄）的研究成果。已有研究发现，儿童社会认知过程中的类别形成离不开社会认知对象所具备的感知觉特征，儿童更容易习得那些具备感知显著性的群体线索。比如，一个月大的新生儿就可以区分与照料者相同性别和种族的其他个体，并在此基础上整合其他方面的信息，逐步形成对于社会群体更为全面的认识。因此，在群体形成过程中，那些在客观物质属性上具有区分性的线索会在儿童社会认知中发挥作用。另一方面，

结合许多研究者的观点，即群体内部的社会义务对儿童来说可能更重要，可以推测，在儿童发展过程中很可能存在着从“感知觉关注”到“社会义务关注”的转变阶段。而以往有关儿童如何理解社会义务这一问题的研究多集中于 4-6 岁的年龄段，而本研究所涵盖的儿童年龄范围更广，为 3-8 岁，因此我们预测较高年龄段的儿童对社会性线索的偏好比较低年龄段的儿童更强。

最后，从认知机制的视角，有关儿童对群体社会性认知的研究尚在起步阶段，近年来仅有 Rhodes 等人的团队专注于这一领域，并尝试从感知基础、构念基础以及社会基础三个方面初步总结了影响儿童分类线索习得和表现的原因，认为三者共同决定了儿童社会认知中的社会分类与加工过程 (Rhodes & Baron, 2019)。对比本研究所提出的群体线索框架，感知基础对应于物质性群体线索，构念基础则对应了社会性群体线索，但依然尚未有研究总结儿童群体认知背后更为深层的认知发展机制，或者讨论其他认知因素的关联。因此本研究也是为解决认知机制方面的问题提供了概念框架、方法和范式上的借鉴，为未来进一步研究儿童群体认知这一主题提供了实证研究的基础。

正文的相关讨论见 1.3 儿童群体认知线索偏好及其发展这一部分。

**意见 2:** 研究的实施采用了线上和线下两种方式，两种测试方式的结果是否存在显著差异？线上的问卷作答是如何进行的？是研究者还是父母代为施测？从论文剔除了作答时间在 300s 以内问卷的表述来看，似乎并非是研究者施测。如果不是研究者施测的话，如何保证被试准确地理解研究的内容并作出相应的反应？线下施测是一一对一进行的还是以集体的方式进行的？线下施测的时长大约是多少？

**回应:** 非常感谢您的具体建议，回答您的问题帮助我们进一步完善了研究。

经过对比线上和线下的数据结果，发现不同的测量方法之间确实存在差异，在三元分类任务中，测量方法的效应显著， $F(1,212) = 9.81, p = .002, \eta_p^2 = .044$ 。进一步分析发现，三元分类任务的测量方法差异主要存在于 3-4 岁的儿童中， $t(92) = 3.39, p = .001, \text{Cohen's } d = 0.90$ 。线下测试的 3-4 岁儿童并不存在对于线索的显著偏好 ( $M = 0.52, SD = 0.34$ )， $t(31) = 0.25, p = .805, \text{Cohen's } d = 0.04$ ，而在线上问卷中儿童则是更为偏爱社会性群体线索 ( $M = 0.71, SD = 0.22$ )， $t(61) = 7.63, p < .001, \text{Cohen's } d = 0.97$ 。

排除任务中同样存在测量方法的效应， $F(1,207) = 9.69, p = .002, \eta_p^2 = .045$ 。进一步分析同样发现仅 3-4 岁的儿童出现了线上和线下测量方法的差异， $t(89) = 3.74, p < .001, \text{Cohen's } d = 0.80$ 。即 3-4 岁儿童在线下并没有表现出线索偏好 ( $M = 0.71, SD = 0.22$ )， $t(27) = 1.75, p = .092, \text{Cohen's } d = 0.33$ ，而在线上存在对于社会性群体线索的偏好 ( $M = 0.71, SD = 0.22$ )， $t(62) = 4.28, p < .001, \text{Cohen's } d = 0.54$ 。

对于上述差异，我们认为可能是儿童在线下作答时，由于施测者为陌生人，以及每次都是由施测者和儿童单独施测的环境，导致儿童出现了紧张情绪，难以全面理解社会性线索的含义。而物质性线索在本研究的呈现方式上更为直观，导致了相对更多儿童在线下任务中选择了物质性线索。然而需要注意的是，尽管线上线下的 3-4 岁儿童的线索偏好上存在差异，但这种差异并未影响到本研究有关儿童社会性线索偏好这一问题的回答，即 3-4 岁儿童已经表现出了对于社会性群体线索的偏好。只是线下数据的分析结果中，儿童在更少的条件下表现出对社会性线索偏好。

考虑到研究方法的差异并未影响到研究的主要结论,而线上和线下测量的差异性确实存在,又因为疫情背景下难以再在线下进行儿童实验,因此为了系统性地考察儿童社会性线索偏好的年龄发展,我们在已有线上数据的基础上补充了 5 名 3-4 岁以及 37 名 5-6 岁儿童以达到研究预先计算的 211 名被试量,所使用的施测工具与之前一致。数据分析显示,新招募的儿童被试在作答结果上与第一批收集的线上数据之间并没有显著差异,三元分类任务中, $F(1,212) = 1.70, p = .194, \eta_p^2 = .008$ ,排除任务中, $F(1,212) = 1.01, p = .316, \eta_p^2 = .005$ 。因此,当前研究 1 共包括有效被试 215 人,具体年龄和性别情况见正文表 1 (同见下表)。

表 1 研究 1 被试性别及年龄分布情况

年龄段 (范围)	性别	<i>N</i>	<i>M<sub>age</sub></i>	<i>SD</i>
3-4 岁 (3.33-4.96)	男生	35	4.01	0.49
	女生	33	4.15	0.42
5-6 岁 (5.02-6.95)	男生	37	6.06	0.60
	女生	39	5.91	0.61
7-8 岁 (7.01-8.40)	男生	36	7.74	0.37
	女生	35	7.64	0.39

线上的问卷作答由父母代为施测。为了保证父母施测的有效性,我们在正式发放问卷前邀请了 3 名 3-6 岁的儿童家长审阅了问卷,并对语言进行修改,以确保家长可以更好地理解问卷主题,保证施测顺利。此外,在正式的问卷中,我们从家长的角度细化了指导语内容,分为作答项目说明以及需要儿童进行反应的具体项目指导语等。其中作答项目说明主要向家长阐明了每部分测量需要家长完成的任务。

在材料学习阶段对应的作答项目说明为:“在问卷中,我们会为您呈现一些小朋友的卡通图片和标志,以下将对卡通人物和标志进行简单介绍,并请您结合图画内容,向小朋友介绍每个人物的特征。”

在三元分类任务之前的作答项目说明为:“接下来我们每次会呈现三个不同的小朋友,每个小朋友身上会有区别于其他小朋友的两种特征,需要儿童从这三人中选择两人进行分类(找到更适合分在一起的两个人)。这部分的回答没有正确和错误答案,需要儿童在辨认出每个小朋友身上所具有特征的基础上,按照自己的想法选择即可。家长需引导小朋友理解题目,感谢您的帮助!”

在排除任务前的作答项目说明为:“接下来这部分我们每次会出现一排四个小朋友,有一个小朋友本来属于别的队伍,但是他站错了,需要儿童选出他认为站错队伍的那个人,选择对应的编号作答。同样需要注意,这部分的回答没有正确和错误答案,需要儿童在辨认出每个小朋友身上所具有特征的基础上,按照自己的想法选择即可。家长需引导小朋友理解题目,感谢您的帮助!”

另一方面,对每个需要儿童作答的项目也都提供了具体项目指导语,举例如下。

学习阶段的具体项目指导语内容如下。性别线索:“这里有两个小朋友,你能认出哪个是男孩,哪个是女孩么?”颜色线索:“你能认出他们的衣服分别是什么颜色么?”经济社

会地位线索：“他们的家分别是什么样的？”群体规范线索：“他们在课堂上会有什么样的表现？”群体归属线索：“他们分别来自哪个班？”兴趣线索：“他们分别喜欢玩的是什么球？”

三元分类任务的具体项目指导语内容如：“这里有三个不同的小女孩，①号穿着白衣服、上课不和别人讲话，②号穿着黄衣服、上课不乱蹦乱跳，③号穿着白衣服、上课不乱蹦乱跳。你是更愿意把③号小朋友和①号分在一起，还是②号分在一起？”

排除任务的具体项目指导语内容如：“这里有一队四个小女孩，他们的衣服颜色和上课时候的表现不太一样，现在有一个小朋友站错队了，你觉得是谁站错队了？”

对于上述列出的所有指导语类型及内容示例均在正文中进行了修改，详见 2.1.1(4) 实验流程和 2.2.1(2) 实验材料与流程。

最后，由于研究 1 已全部为线上研究，因此仅有研究 2 为线下研究。在研究 2 中，所有儿童均为一对一施测，完成研究 2 大约需要 10 分钟左右。这一信息已补充在正文研究 2 的研究流程中（详见 3.1.4 研究流程）。

**意见 3:** 2.2.2 部分，“幼儿有可能会依据位置和顺序等原因选择参照任务，因此将此类选择记为缺失值”，这句话的表述比较模糊，请举例说明。另外，因此产生的缺失值比例是多少？这些缺失值在统计分析中是如何处理的？

回应：感谢您的建议。这里主要是因为在线下预实验以及数据分析中都发现，儿童在选择四人中的哪一个应该排除的时候，偶尔会选择非研究者预设的目标人物，即在物质性或社会性群体线索上与其他三个人不同的人物的图片。因为儿童在选择的时候可能会有他们自己的逻辑，比如他们可能会说因为这个人排第一个，其他小朋友也想排第一个，所以按照顺序排除了第一个小朋友等等，而不是因为群体线索的一致性而选择特定人物。

另一方面，尽管存在这种情况，但在数据预处理过程中我们发现，在 214 名儿童中，仅有 20 位儿童出现了选择参照人物的情况，对应的选择次数在 1-5 次之间，而当分析所有儿童在九种实验条件下的反应，发现在 1926 个 trial 中，仅有 41 个 trial 为这种情况。因此这些 trial 被转化为了缺失值。考虑到本研究的目的主要是了解儿童在物质性线索和社会性线索之间的偏好，而此类反应并不能够帮助回答研究问题，因此在后续分析中此类 trial 均被删除。根据您的建议，我们在修改时进行了补充说明，详见 2.2.1(3) 数据分析方法。

**意见 4:** 质性研究样本数量确定的依据是什么？为什么 3-4 岁组和 5-6 岁组的人数差异比较大？3-4 岁儿童同时呈现八张图片，并要求其进行分类，这一任务的难度是否对这个年龄段的儿童来说太大？这个年龄段的儿童通常很难同时注意和记忆这么多的信息。

回应：感谢您的建议。在本研究中，质性样本的数量确定主要是依据经验值，参考了质性分析实现编码饱和（12 个样本量可提取 92% 的编码）和意义饱和（16-24 个样本）的样本要求（杨莉萍 等, 2022），取中间值（15 人）确定了样本量大小。然而，由于疫情因素，5-6 岁的样本量并未达到预期，但依然大于实现编码饱和所需要的样本数量。相关内容在正文中也进行了补充。

之所以本研究中两个年龄段被试样本量差异较大，是因为研究 2 的数据是在疫情爆发前，即 2019 年 12 月底到 2019 年 1 月初采集的，当初的计划是从幼儿园小班开始实验一直到大班，然而后来受到疫情影响，难以到幼儿园进行实地的调查研究，因此质性研究样本量没有完全达到计划样本量，5-6 岁儿童人数较少。但由于在研究设计中，每位被试都需要完

成两个不同条件下的分类任务，包括三次八人分类和一次五人分类，因此 15 名 5-6 岁的儿童依然提供了相当数量的有效分类行为（99 次）和原因说明（87 条），与 3-4 岁儿童（分类行为 123 次，原因说明 113 条）的差异并不是很大。因此，总体来说，研究二对于 3-4 岁和 5-6 岁的分类偏好对比依然是有意义的。

关于研究 2 中的八人分类任务对较低年龄段儿童的适用性问题，在实际的操作中我们发现，确实存在部分 3-4 岁儿童无法很好地理解实验目的并做出分类行为（6 人），这部分儿童也在后续分析中被完全排除。而对于余下的 29 名 3-4 岁儿童来说，他们事实上首先完成了三元分类和排除任务的预实验，已经熟悉并了解了研究材料，并在适当休息后再开始研究 2 的具体任务。其次，在进行研究 2 的操作前，我们也为儿童安排了再认环节，即在所呈现的 8 张图片中，根据主试的指导语找出对应的线索类别，在完成了以上流程后再正式开始群体分类任务。最后，可能对于部分儿童来说，八人分类确实有一定的难度，因此我们也在后面进行了一次五人分类的任务，并分析了五人分类中儿童所表现出的线索偏好，其结果与八人分类结果间具有一致性，说明在当前研究中，3-4 岁儿童有能力理解和加工每次为他们所呈现的群体线索。

总之，研究中所保留下来的 3-4 岁儿童均在八人或五人分类任务中表现出了有意义的分类行为，或提供了有效的分类原因。经过上述流程，我们认为 3-4 岁儿童在群体分类中所提供的数据和结果是有效的。

**意见 5:** 质性分析部分，大部分结果中的比例都很低，且有可能受到极端值的影响。这些比例之间的差异是否显著？这些比例之间的比较是否有意义？

**回应:** 非常感谢您的建议，的确大部分结果的比例都比较低。为了解决这一问题，我们查阅了以往涉及儿童分类行为的质性研究，并参考 Burkholder, Elenbaas, 和 Killen(2021)的资料整理方式，对研究 2 中儿童的行为和原因分开进行了重新整理和编码，最后获得了 222 次有效分类行为和 200 条有效原因说明。在呈现编码结果时，同时呈现了编码所对应的行为或原因数量占总体有效行为或原因数量的比例。这种方式一方面提高了编码结果的效度，同时也使得年龄之间的比较在量化层面上更有意义。

当前论文也对原因编码类型进行了进一步改进。考虑到儿童首次分类的线索选择也体现了他们对线索的偏好，因此当前研究将儿童首次分类中所使用的群体线索从质性的原因编码中独立出来，形成了量化分析部分的新的因变量（结果详见 3.2.1 首次分类的群体线索偏好）。而在儿童群体分类原因的质性分析中，在群体线索这一类别编码下，更关注儿童在使用群体线索分类时所表现出的语言特点，分为了“具体特征”和“线索抽象”两大类（结果详见 3.2.3 儿童群体分类的原因归纳）。

我们认为这种改进一方面更直观地用数据回答了不同年龄段儿童是否具有社会性线索偏好这一问题，另一方面也使得质性分析中涵盖了更多与儿童发展有关的信息，有助于启发后续有关儿童语言发展和群体分类能力之间的研究。

**意见 6:** 讨论部分，重点还是应该放在对发展特点背后机制的分析上。研究的理论价值和应用价值可适当缩减。此外，图 6 的结果从何而来？对应研究中的哪部分结果？0-3 岁以及 8 岁以上的发展特点是基于什么提出的？

**回应:** 非常感谢您的建议，在您和另一位审稿人意见的基础上，我们结合新补充的研究结果，

重新构思并修改了整篇讨论的内容，重点讨论了①儿童对于社会性线索的偏好；②儿童社会性线索偏好的发展及可能机制；③物质性与社会性的群体线索框架。其中，第二部分对儿童社会性线索偏好的发展过程及其机制进行了较为细致的探讨，详见正文 4.2 群体分类中社会性线索偏好的发展及可能机制。

至于研究的理论与应用价值的部分，结合两位审稿人的建议，同时也考虑到有关儿童物质性和社会性线索的研究尚在起步阶段，缺乏实证研究的支持，因此在当前论文中删去了线索偏好发展趋势的模拟图。此外，为了避免过度推论，本研究同时也精简了相关内容，以物质性和社会性线索这一新的线索类型框架为核心，讨论了其内涵和研究价值，并指出了未来可能的研究方向。详见正文 4.3 物质性与社会性的群体线索新框架。

**意见 7：**结论部分的表述需更加精确，本研究中真正的 3 岁儿童可能并不多，因此结论讲 3 岁儿童表现出对群体本质线索的偏好不准确。此外，3-8 岁儿童对群体本质线索的偏好逐渐增加，这一表述也不精确。这一阶段儿童的偏好发展也有可能出现阶段性的变化，而非量的积累。本研究无法排除这一可能。

**回应：**非常感谢您的建议，我们更新了表述。在新招募了研究一的线上数据后，本研究中明确自我报告年龄在 3 岁（3.33-3.99）的儿童数量为 30 名，占总被试数量的 14.01%，与 1/6（16.67%）较为相近。我们额外分析了 3 岁儿童总体的群体线索偏好发现，在三元分类任务中，3 岁儿童（ $M = 0.71, SD = 0.19$ ）存在对于社会性群体线索的偏好， $t(29) = 6.12, p < .001$ ，Cohen's  $d = 1.12$ ，在排除任务中，3 岁儿童的社会性线索偏好（ $M = 0.61, SD = 0.31$ ）为边缘显著， $t(29) = 1.99, p = .056$ ，Cohen's  $d = 0.36$ 。在研究二中，3 岁儿童数量确实较少，仅为 4 人，且 3-4 岁儿童在分类线索选择上更多选择了物质性群体线索，在本研究中为性别和经济社会地位。对于这一结果，我们认为可能是研究方法本身的复杂性和这两个线索在儿童生活中的熟悉性共同导致的（详见讨论 4.1 儿童在群体分类中的社会性线索偏好）。

此外，对于本文中“偏好逐渐增加”的表述确实不太准确，正如您所说，本研究中是将年龄视为分类变量，分别讨论了不同年龄段儿童的群体线索类型偏好，因此在相关的结果表述中，我们将相关描述改为“较高年龄段的儿童相比较低年龄段的儿童更偏好社会性线索”。以 5 结论部分为例，具体内容如下。

- (1) 3-8 岁儿童能够在群体分类中认知和识别物质性线索和社会性线索。
- (2) 3-8 岁儿童在群体分类中总体呈现出对社会性线索的偏好。
- (3) 高年龄段儿童比低年龄段儿童的社会性线索偏好更明显，并且在不同情境间更为稳定。

**意见 8：**术语的英文标注在文中只需出现一次就好，不需要重复出现。

**回应：**非常感谢您的建议。我们已经检查了论文当前版本的英文标注，以确保没有重复。

**意见 9：**英文题目与中文题目的对应不佳。

**回应：**感谢您的建议，已根据研究的假设与内容，修改了论文的中文题目并修改了英文题目。目前中文标题为“3-8 岁儿童在群体分类中的社会性线索偏好：实验验证与质化研究”；英文标题为“The preference for societal cues in 3- to 8- year-old children's group categorization: Experimental and qualitative evidence”。

.....

## 审稿人 2 意见：

我认为本篇文章会有潜在的理论意义，提出的问题也没有和已有研究雷同。但是文章目前还存在比较大的理论问题，主要是关于“群体视觉线索”和“群体本质线索”的内涵界定和具体操作。如果不加以修改，那么目前采用的方法和结果就没能回答作者提出的问题。我提出了一个比较可行的修改建议，作者可以考虑。具体意见见附件。

### 【附件内容】

本篇文章研究“群体视觉线索”和“群体本质线索”对儿童群体分类的影响，发现 3-8 岁儿童对于群体本质线索的偏好强于对群体视觉线索的偏好，而且这个倾向随着年龄增长变得更强。该研究在理论和方法上有一定的创新性，结果也对于领域发展有一定的积极影响。以下是我关于文章的一些主要方面的问题和想法，希望对于作者改进文章有所帮助：

### 【主要理论问题】

作者提出了两种群体线索的类型，“群体可视线索”（能够通过视觉信息区分的群体线索）和“群体本质线索”（预示了群体规范和内部成员关系）。作者认为可视线索对应某些群体类别（颜色，性别，地位），抽象线索对应某些群体类别（爱好，归属感，行为规范），但是这种划分存在比较大的理论问题。

根据作者对两种线索的定义，逻辑上来讲每个具体的群体类别都应该可以通过两种线索进行体现和划分。以社会经济地位为例，富人和穷人确实可以通过明显的视觉信息加以具像化和区分（比如住的房子），但是还有更本质的信息是不可视的，和软实力及社会关系有关的（个人认同感，教育文化程度，社会关系网络等）。所以认为“社会经济地位”代表了“群体视觉线索”是有问题的。作者甚至在文中都提到了社会经济地位“也是由社会塑造的群体线索”（p.9）。试想，如果没有呈现不同类型的房子图片给儿童，而只是描述“这个人很富”“这个人很穷”，那么这个操控的还是“群体可视线索”吗？性别也会存在类似的问题。如果呈现两个长得一样的儿童图片，但是介绍说“这是个男孩，这是个女孩”，那么操控的还是“可视线索”吗？

如果想研究“群体可视线索”和“群体本质线索”，最直接的方式应该是研究对于同一种类别（比如社会经济地位），当两种线索冲突时儿童如何分类（比如一个有软实力的，却表面上看起来的生活俭朴的人是什么社会经济地位），而不是把社会经济地位和行为规范或者归属感对立。根据研究结果，儿童对于这种对立条件的回答和其他的条件很不一致，从某种程度上也说明了目前的理论划分存在问题。

所以，基于作者目前的实验设计，我认为得到的结果和结论不能支持关于“群体可视线索”和“群体本质线索”的儿童偏好的结论。作者需要考虑整体改变对研究问题，导论和结论部分的陈述。

一种可能的方法是重新定义两个大的理论类别，并且对于分类方式需要提供更多的理论解释和讨论。“爱好，归属感，行为规范”的共同点是都有社会性，这一点和作者目前的界定是比较接近的，但是“颜色，性别，地位”的共同点不是他们都是“可视线索”，而是可以概括为它们都有客观性（不完全是由社会标准和主观感受决定的）。所以问题和结论都可以概括为关于“社会性群体线索”和“客观性群体线索”的偏好。这样的分类仍然有一定的人为性（如果看成一个连续变量而不是两分的维度会更合理），但是会比目前的设定更准确一些。

**意见 1:** 如果作者不修改目前对问题, 导论和结论的陈述, 那么目前的方法就存在比较大的问题。应该考虑加一个实验, 用完全不同的方法来更严谨的研究提出的问题, 比如对于同一类别, 看当两种线索冲突时儿童如何分类。如果作者对整体陈述进行修改, 那么目前的方法大致上没有严重的问题。

**回应:** 非常感谢您的清晰且创新的建议, 线索名称确实是核心问题, 也是我们的创新点。由于在当前疫情的背景下, 很难在幼儿园开展新的研究, 因此我们根据您的提议, 反复研读相关文献, 最终决定按照您的思路修改了群体线索类型的名称, 将群体线索分类由原来的群体“可视线索”和“本质线索”改变为群体“物质性线索”和“社会性线索”。同时, 论文中对于相关的理论和实证研究基础也进行了重新整理和修改, 详见下文(斜体部分)。

## 1 前言

### 1.1 儿童群体分类线索

儿童对于群体线索的认知与识别能力发展很早, 6个月左右的婴儿就已经能够根据性别或种族线索区分他人(如 Martin, Ruble, & Szkrybalo, 2002; Kelly et al., 2007)。总结以往儿童研究所关注的群体线索可以发现, 大多数都集中于西方文化背景下“大三”的社会分类线索, 即性别、种族和年龄(Weisman, Johnson, & Shutts, 2015; Lei, Leshin, & Rhodes, 2020), 仅少数研究关注了儿童对于职业(Bian, Leslie, Cimpian, 2017)、地位(Shutts, 2015)等群体线索的感知与分类。然而, 在这些既定的社会群体基础上进行的群体研究具有局限性, 因为群体概念本身所强调的是成员间的紧密联系和一体感。群体中的成员往往在心理上相互依赖, 行为上存在互相影响(林崇德, 杨治良, 黄希庭, 2004; Turner, 1982)。在此基础上, 越来越多研究者开始关注儿童对群体内社会关系的理解(如 Roberts, Gelman, & Ho, 2017; Switzer et al., 2020)。

当前的儿童研究中, 对群体内社会关系的讨论集中于社会义务(social obligation; Rhodes, 2013)和群体规范两个方面, 而群体分类线索在其中都有所体现。比如在探索儿童对社会义务的理解时, 研究者往往使用群体成员的共同偏好或共享目标区分群体, 在此基础上关注儿童对群体内或群际关系的预测结果(Jordan & Dunham, 2020; Roberts et al., 2017; Switzer et al., 2020)。而在群体规范的讨论中, 有研究在性别群体基础上讨论了具体特质和行为对评价的影响(Bian et al., 2017; 牡丹, 苏彦捷, 2005); 也有研究发现成员观点与群体观点之间的一致性也会影响到儿童对特定成员的评价(Roberts et al., 2020), 即成员观点也可以成为区分群体的线索之一。由此可见, 儿童能够理解的群体线索并不限于与生俱来的线索(如性别、种族), 也不只是社会角色所塑造的线索(如职业、地位), 还包括了能够反映成员间社会关系的线索(如行为、爱好)。

### 1.2 物质性和社会性(群体)线索

群体线索可以根据自身特征被归入不同的类型或维度, 有研究者从不同的视角进行了划分, 如自然线索和社会线索, 明显线索和模糊线索, 静态线索和动态线索等(温芳芳, 佐斌, 2019; 佐斌, 温芳芳, 宋静静, 代涛涛, 2019)。然而, 这些分类多是对已有研究所涉及群体的归纳, 缺乏发展视角和相应的理论阐述。而 Rhodes(2013)等人提出的社会分类的直觉理论(intuitive theory of social categorization)为儿童研究中群体线索的类型框架提供了参考。

直觉理论认为儿童对群体的理解可能存在两类先天的认知倾向: 其一是将群体看作是自然类别(natural kinds), 儿童会将群体线索视为先天且稳定的结构, 线索划分的群体之间(如男性和女性)截然不同, 群体内部相似性更高; 其二是将群体与社会义务挂钩, 群体内成员



更可能产生良性互动并为彼此提供社会支持(Rhodes & Chalik, 2013)。这一框架与群体实体性理论中本质论与主体性视角(Brewer, Hong, & Li, 2004)或类别性和动力性视角(Rutchick, Hamilton, & Sack, 2008)之间具有一致性: 自然类别、本质论和类别性都强调群体内成员在物质属性上的相似, 而社会义务、主体性和动力性则关注群体内成员间的社会关系和互依模式(杨晓莉, 刘力, 李琼, 弯美娜, 2012)。

总结上述理论可以发现, 人们对群体的理解存在两个不同方面, 分别为群体成员在外表特征或物质属性上的一致性和群体内成员间的社会性联结。然而, 以往研究对这两方面的操纵并不等同: 研究者多以评价对象在外表物质特征上的相似或差异定义群体, 同时又使用不同的行为模式操纵群体内成员间的社会关系(如 Bian, Sloane, & Baillargeon, 2018; Switzer et al., 2020)。这种比较忽视了由外表相似性定义的群体和行为模式本身在人们的社会认知过程中的权重差异, 尤其当我们把行为模式理解为个性化(individuation)信息的时候(相关讨论见 Fiske & Neuberg, 1990; Monroe, et al., 2018)。因此, 有必要需要统一二者的呈现方式, 比如都以线索的形式体现群体的外在物质属性与内部社会联结特征, 以排除其他因素在其中的影响。

在相关理论基础, 本研究采用新的维度划分了群体线索类型, 即物质性(physical-type cues)和社会性(societal-type cues)线索。具体而言, 物质性线索主要指由成员所表现或拥有的物质或物理特征所定义的群体线索。其中, 既包括了与生理特征息息相关的“大三”线索(如 Fiske, Lin, Neuberg, Chaiken, & Trope, 1999), 也包含了由人所附属的物质属性定义的线索, 如衣服颜色(Chalik & Rhodes, 2014), 以及在社会范围内由人所占有的物质多寡而定义的社会经济地位等。社会性线索则是能够反映或表现群体内部成员关系的线索, 体现了成员对于群体的积极认同。因此, 从群体线索的视角, 群体归属(Platow, Grace, & Smithson, 2012)可以被视为群体认同的直接表现形式, 而以往研究所关注的行为规范(Foster-Hanson & Rhodes, 2019; Roberts et al., 2020)和共同爱好(Jordan & Dunham, 2020; Switzer et al., 2020), 都体现了成员对于群体在行为方面要求的自发遵守。

**意见 2:** 实验 1 当中的因变量是问“哪一个小朋友会更愿意和另外两个小朋友成为朋友”, 这是一个常用的关于社会认知的推测性问题, 但是准确来说个问题测的是儿童作为第三方对于他人的社会关系的推断, 而不是代表儿童自身对于群体分类的偏好。实验 2 的变量相比之下就更加准确。作者应该对此不足加以讨论。

**回应:** 非常感谢您的建议。指导语的不同确实可能会对结果及其解释产生影响, 但在实验 1 中, “交友”的指导语实际上是在儿童无法理解“分类”指导语的情况下进行的补充说明, 在实际实验过程中出现的次数很少。

在正式实验开始前, 我们选取了幼儿园小班和中班的部分儿童对实验 1 和实验 2 的研究流程进行了预测验, 在实验 1 的三元分类任务中, 主试会首先询问儿童: “你是更愿意把他们俩(对应分类线索 1)还是他们俩(对应分类线索 2)分在一起?” 在这个过程中会借由指示的肢体动作向儿童说明。但与此同时我们发现, 有部分 3-4 岁儿童并不能理解这个指导语的意思, 做出有效的分类反映。因此, 我们会再次询问儿童: “你觉得是他们俩(分类线索 1)更愿意做朋友, 还是他们俩(分类线索 2)更愿意做朋友?” 在这种情况下, 几乎所有的儿童都可以对实验任务做出反应。在之前的论文写作中, 由于在实验流程介绍时主要以

线下实验为例介绍了研究的主要过程，为了尽量精简写作，使内容易于理解，我们对指导语的内容进行了少量删减，忽视了“分类”和“交友”的指导语差异可能带来的影响，使实验 1 的指导语看起来并不能反映群体分类偏好这一变量。

在当前版本中，由于线上和线下儿童在线索偏好上的差异，所以研究 1 只使用了线上研究的数据，并根据已招募年龄段的儿童数量补充了儿童被试，因此当前正文的指导语内容修改为线上版本，整个施测在家长的帮助下进行。为了保证指导语的准确性，我们为家长呈现了需要他们自己阅读的作答项目说明和可参考的具体项目指导语，并在正式发放问卷前邀请了 3 位幼儿园儿童家长对指导语内容进行了阅读和修改，其中与三元分类任务直接相关的两类指导语内容示例如下。

作答项目说明为：“接下来我们每次会呈现三个不同的小朋友，每个小朋友身上会有区别于其他小朋友的两种特征，需要儿童从这三人中选择两人进行分类（找到更适合分在一起的两个人）。这部分的答案没有正确和错误答案，需要儿童在辨认出每个小朋友身上所具有特征的基础上，按照自己的想法选择即可。家长需引导小朋友理解题目，感谢您的帮助！”

具体项目指导语内容如：“这里有三个不同的小女孩，①号穿着白衣服、上课不和别人讲话，②号穿着黄衣服、上课不乱蹦乱跳，③号穿着白衣服、上课不乱蹦乱跳。你是更愿意把③号小朋友和①号分在一起，还是②号分在一起？”

结合以上内容我们认为，在实践中实验 1 的指导语依然同群体分类这一主题紧密相关，其结果与实验 2 结果相互印证，共同回答了儿童在群体分类过程中的线索偏好问题。

**意见 3：**作者需要加以说明样本量是基于 GPower 当中哪种具体的分析计算出来的。

**回应：**非常感谢您的建议，当前的论文中已经对于样本量计算方法进行了详细说明：“研究将年龄组和儿童性别视为两个独立的自变量，使用 G\*power 软件中的 F 检验计算了研究所需的样本量，设定效应量为 0.25，显著性水平为 0.05，发现达到 0.80 的统计检验力需要 158 人。”

**意见 4：**关于每种条件的操控的具体的指导语（颜色，规范，归属，兴趣……）属于关键性实验信息，建议至少在实验 1 中的正文中（而不是补充材料中）加以描述。

**回应：**非常感谢您的建议。结合另一位审稿人的建议，当前的研究 1 在线上补充了部分儿童被试使之满足了 211 人的被试量需要后，仅分析了线上的数据。因此，正文中对应补充了需要家长阅读的作答项目说明和需要儿童回应的指导语两部分，并以举例和附注的形式对指导语内容进行了详细说明，详见 2.1.1（4）实验流程。

最后要说明的：我们依据评审专家的意见和建议，反复思考，补充研读文献，理清楚我们的研究问题、假设以及研究的操作。依据研究的假设和内容，修改了论文题目，进一步检查完善了语言表达、标点符号、图表格式、参考文献以及中英文摘要。所有的修改在正文中都以红色字体进行了标记。对本文所进行的较大修改简要列举如下：

1. 在分析和比较线上和线下儿童数据结果基础上，同时考虑到国内疫情防控的要求，对研究 1 所涉及的三元分类和排除任务我们重新补充了儿童的线上数据，采用了和之前线上数据收集相同的方法，使研究 1 中的样本量达到了 G\*power 计算得到的 158 人的要求。

2. 结合两位审稿专家的意见，在查阅文献和反复论证、思考的基础上，我们修改了文

章重点关注的两类线索的命名，由最开始的可视 vs. 本质的群体线索，改为了现在的物质性和社会性群体线索，并在文献综述部分给出了线索定义，阐述了其背后的理论基础。

3. 在文献综述和讨论的部分，我们重新构思了相关部分的写作逻辑，增加了对儿童社会性线索偏好及其发展的理论和实证研究基础的说明，删减了讨论部分对于儿童社会性线索偏好发展趋势的过度推论，突出了其背后可能存在的认知加工机制。

总的来说，此次修改不仅补充了实验数据，更重要的是在评审专家的帮助下，进一步理清了儿童对群体知觉时的分类线索的理论框架和名称。本次修改我们用了很长的时间，力图提高本研究的质量。衷心感谢评审专家的意见和建议，我们论文如果能够得到认可，离不开评审专家的宝贵建议！

---

## 第二轮

**审稿人 1 意见：**作者针对审稿意见进行了非常认真的修改，论文整体有一定的提升。

**意见 1：**儿童社会分类加工的发展特点是研究的关键，修改稿虽然补充了相关的参考文献，但对发展特点假设的论证仍然非常弱。作者指出，“考虑到物质性线索相关的知觉显著性对儿童社会分类的基础性作用(Rhodes & Baron, 2019)，以及相比物质性线索，儿童对社会性线索的关注开始更晚(Liu et al., 2015; Pun et al., 2021)。本研究认为，随着年龄增长，儿童更可能会表现出依赖社会性线索分类和认知群体的倾向”。问题是这与本研究关注的 3-8 岁年龄段的关系是什么？从前面两点，如何能推断本研究所关注群体的发展特点？

**回应：**非常感谢您的建议。论文之前的综述部分确实忽视了本研究所关注的 3-8 岁年龄段，花费了过多篇幅论证物质性和社会性线索的来源。结合您的建议，我们在新修改综述的 1.3 部分论述了：①为什么本研究关注了儿童群体认知线索偏好这一问题；②以往研究关于 3-8 岁儿童的群体线索偏好问题有什么样的研究发现。其中有关第二点的文献综述围绕 5-6 岁这一儿童群体认知发展关键时期展开。

以往研究发现儿童在 5 岁左右开始应用社会对于群体的期望，从心理或行为特征的视角理解群体身份对个人的意义。此时，他人行为是否遵守社会或者群体规范才是儿童群体认知与评价时的考虑因素。也就是说，在这一时期，儿童对于群体的理解才可能涉及到“社会关系”或者“个体认同”的层面。基于此，我们推测对于本研究中所关注的社会性线索偏好可能也存在着类似的发展趋势，将 3-8 岁定为本研究儿童被试的年龄范围，并假设 5-6 岁是儿童社会性线索偏好发展的关键时期。这一范围的选择一方面与以往研究提出的 5 岁左右群体认知发展关键期相互呼应，另一方面也便于区分三个年龄段以进行发展时序上的比较。

对于这一问题的详细修改见正文 1.3 儿童群体认知的线索偏好及其发展的相关研究，以及 1.4 当前研究 中的部分文本（已用红色文字标注）。

**意见 2：**作者在修改稿的研究一里排除了线下数据，补充了线上数据，并基于线上数据进行了分析。结果发现，初稿中的显著年龄差异变得不再显著。研究的整体结论也变为 3-8 岁儿童均能在群体分类中识别物质性线索和社会性线索，并呈现对社会性线索的偏好。这是一个

相对笼统的结果，难以对 3-8 岁期间儿童群体分类加工特点进行精细的描述。当然，这背后有两种可能的原因，一是 3-8 岁期间儿童的群体分类加工特点的确比较稳定，而另外一种可能则是本研究的方法并未准确反映可能存在的年龄差异。线上作答存在的一个可能的弊端是研究者对实验操作的控制性大大降低。研究无法排除父母在进行测试时是否对孩子的作答进行了引导。从初稿和修改稿的原始数据来看，线上作答时，3-4 和 5-6 岁组的得分有一定程度的提高。这是否有可能与父母的引导有关？作者指出，线下作答会导致儿童出现紧张情绪，难以理解社会性线索的含义。首先，紧张情绪是否只会影响低年龄段儿童的作答？其次，在进行幼儿相关研究中，研究者往往都会采取一些措施，降低儿童的紧张程度。这是一个可以控制的变量。

**回应：**非常感谢您的建议。本研究在线上 and 线下结果间的差异也让我们非常纠结，尤其是在以何种形式呈现报告结果这一问题上。相比之下，我们更加信任线下实验而非线上实验的结果。因为线下实验得到了幼儿园老师和家长的支持，在儿童熟悉的幼儿园环境中完成施测，在实验前进行了周密的安排和计划，开展了主试培训，最终顺利完成施测。而线上实验尽管采取了一系列措施以排除家长参与的影响，但由于研究平台的限制，必须家长辅助施测，而家长可能存在的干扰并不能通过实验设计完全排除。

通过查阅文献，我们发现研究对较低年龄段儿童线上和线下施测方式的影响进行了检验，发现 3-4 岁儿童线上施测的结果更加不可信，而 5-6 岁儿童线上线下测验的结果相对较一致(Lapidow, Tandon, Goddu, & Walker, 2021)。因此，我们引用了这一研究，并重新整理了参与线下实验的 3-4 岁儿童被试的数据。而考虑到线下实验对应了研究 1 的两个实验，以随机顺序先后对儿童进行了三元分类和排除任务的测量，如果以新实验的形式写入研究 1，会导致研究 1 的结构不清晰：前两个实验都在线上测量了 3-8 岁儿童，差异在于实验任务不同；而实验 3 只测量了一个年龄段，包括了实验 1、2 的两种任务，但改变了施测方式。这种实验呈现方式会显得逻辑混乱，影响读者的阅读感受。因此，新修改的版本仅在研究 1 的讨论中概述了线下实验的内容和结果，比较了线上和线下实验结果间的差异，最终决定采信 3-4 岁儿童的线下实验结果，详见 2.4 讨论。

此外，为了实验结果报告的完整性，也期望未来能有研究进一步分析线上和线下施测的变异来源，讨论线上和线下实验结合的实验方式。我们在此也对 3-4 岁儿童线下实验的流程和结果进行了报告（同见附录 2），具体内容如下。

#### 实验方法

线下共招募幼儿园小班到中班儿童共 38 名（3.29-4.99 岁），其中有 3 人未完成三元分类任务，6 人未完成排除任务，最终完成两个任务的详细被试情况见表 A2。

表 A2 线下 3-4 岁儿童被试的年龄情况

性别	三元分类任务			排除任务		
	<i>n</i>	<i>M<sub>age</sub></i>	<i>SD</i>	<i>n</i>	<i>M<sub>age</sub></i>	<i>SD</i>
男	18	4.31	0.39	15	4.40	0.40
女	17	4.24	0.42	17	4.25	0.36

本实验由两位预先培训的女性实验者施测完成。实验所使用的材料及操作流程与研

究 1 的两个实验基本相同。所有测量在儿童较为熟悉的幼儿园环境中完成，实验者对场地进行了事先布置，并在作答前后和中间安排了休息环节，以避免儿童的疲劳和注意力不集中对实验结果造成额外影响。实验过程中的指导语对应线上问卷中需要儿童做出反应的指导语内容。儿童被试在通过线索学习之后以随机顺序完成三元分类和排除任务。

参考研究 1 分析思路，实验对两个任务分别计算了儿童社会性线索偏好得分。预分析发现，与研究 1 结果相似，性别的效应均不显著 ( $p_s > .172$ )。因此正式分析首先使用单样本  $t$  检验比较了儿童在两个分类任务中是否存在显著的社会性线索偏好，之后使用二项检验分析了儿童在具体实验条件下的社会性线索选择比率。在排除任务中，有 17 名儿童选择了参照人物，每人出现该类选择的次数在 1-5 次；在所有 288 个试次中，有 38 个试次因此被记为缺失值 (13.19%)。

### 实验结果

实验首先比较了儿童在两个实验任务中的社会性线索偏好得分和随机水平 (0.5) 的差异，发现在三元分类任务 ( $M = 0.55$ ,  $SD = 0.35$ ) 中，3-4 岁儿童在线下施测中并未表现出社会性线索偏好， $t(34) = 0.81$ ,  $p = .426$ , Cohen's  $d = .14$ ；在排除任务中 ( $M = 0.43$ ,  $SD = 0.26$ )，儿童对两类线索的选择同样近似于随机水平， $t(31) = -1.47$ ,  $p = .153$ , Cohen's  $d = -.26$ 。

随后实验对具体条件下儿童的选择进行了二项检验 (表 A3)。结果发现，在所有条件下，3-4 岁儿童对两类线索的选择比率近似相同，即该年龄段儿童在线下的群体分类任务中并没有表现出对社会性线索的选择偏好。

表 A3 线下 3-4 岁儿童在群体分类任务中的社会性线索偏好比例及二项检验结果

实验条件	三元分类任务		排除任务	
	%	$p$	%	$p$
颜色×规范	42.86	.500	56.00	.690
颜色×归属	60.00	.311	53.85	.845
颜色×兴趣	65.71	.090	39.29	.345
性别×规范	63.64	.163	33.33	.099
性别×归属	65.71	.090	48.15	> .999
性别×兴趣	67.65	.058	48.28	> .999
社会经济地位×规范	34.38	.110	35.71	.185
社会经济地位×归属	34.38	.110	55.56	.701
社会经济地位×兴趣	50.00	> .999	50.00	> .999

最后，我们在总讨论中也对施测方式给研究结果带来的影响进行了总结，并在研究不足和展望中对此进行了反思，详见 4.1 3-8 岁儿童群体认知中的社会性线索偏好，以及 4.4 研究不足与展望 中红字标注内容。

**意见 3:** 作者在修改稿中将核心概念调整命名为“物质性线索”和“社会性线索”，物质性和社会性是否是对应的概念？哲学上来讲，物质性指的是客观实在性，物质往往是和精神相对应

的概念。

回应：非常感谢您的建议。事实上，在上一次修改中评审专家也曾给出有关命名的建议，我们在此基础上反复思索了核心线索类型的命名方式问题。最终，结合理论和具体实证研究内容，我们决定将两类线索命名为“物质性线索”和“社会性线索”，具体原因如下。

首先，本研究是以群体认知中的“社会性”为核心展开的。我们虽然划分了两种线索类型，但更关注的其实是社会性线索。结合以往研究成果，我们预期在个体发展过程中存在着对群体的“社会”认知发展的关键时期。这一观点来源于班杜拉的社会学习理论，相关观点在儿童的性别和群体认知发展研究中被广泛应用。研究者认为，儿童对于群体概念的学习经历了从自主建构到社会学习的过程。在自主建构阶段，儿童从周围环境中获得关于群体外在特征，关联线索等信息，初步习得群体类别；而在社会学习阶段，他们会从周围人的观点和态度中习得规范性信息，包括社会和群体自身对成员的规范性要求，形成对群体成员的态度 (Bigler & Liben, 2007; Martin, Ruble, & Szkrybalo, 2002)。也就是说，在社会学习阶段，儿童开始在意“他人”的观点和看法，用以指导自身的社会认知或群体认知方式。这种对“人”的关注，即本研究所讨论的“社会性”的核心体现。对于社会分类任务来说，“人”或者“社会性”的内涵主要就表现在儿童对能够体现成员群体认同的线索的关注。在本研究中，不论是群体归属、共同兴趣还是群体规范都和成员的认同紧密相关，只有当成员认同群体身份，他们才会和群体中的其他人表现出同样的行为方式。因此，我们将这类线索命名为“社会性线索”。

其次，“物质性线索”的命名与本研究在具体线索的操作性定义有关。除了社会性线索，研究所选择的群体线索分别为性别、衣服颜色和社会经济地位。以往研究中，这三种线索很少被归为一类：性别是最常见的“大三”线索之一，常被认为是自然线索；社会经济地位高低取决于社会比较，常被认为是社会线索 (温芳芳, 佐斌, 2019)；而衣服颜色是一种常见的最简群体划分方式，很少与另外两种线索并列讨论 (温芳芳, 佐斌, 2018)。因此，本研究将三者归为一类本身也是一种创新。一开始，我们将这三类线索的共同点归于可视化，即它们都可以通过直观的、物理的量化指标反映。但受限于研究设计，评审专家认为“可视化线索”的命名方式存在缺陷，并给出了“客观性线索”的建议，与“主观性线索”相对。然而，从我们选择的群体类型来看，性别和社会经济地位也并不“客观”。性别概念背后存在着复杂的性别认同，而社会经济地位则是由人“主观”建构的社会决定的，且个体感知到经济社会地位的高低随着“主观”选择的比较对象改变而改变。因此“客观性”也不能反映此类线索的共同点。论文修改过程中，我们从最初选择这三类线索的根本原因，即他们都可以由直观的物理属性反映这一特点中得到启发，先确定了“physical-type cues”的英文名称，之后选择了“物质性”作为最终命名。之所以没有选择由 physical 直译的“物理性”，是因为中文语境中“物理”脱离了本研究中对于人或群体的关注，而这三类线索都是对人或者与人相关的物质的物理属性的反映。

最后，关于您提到的两类概念是否对应的问题，虽然在马克思主义哲学中常把“物质”和“精神”视为一对概念，但马克思主义研究者科亨也指出，在马克思主义的历史唯物主义理论中，同样隐含着“社会的物质性”和“社会的社会性”这一对概念。虽然借助这对概念理解历史唯物主义观点的方式受到了部分研究者的批判 (孟庆龙, 2011)，但这种“物质性”和“社会性”划分方式依然可以为本研究提供借鉴。

综上所述，我们的论文最终保留了“物质性线索”和“社会性线索”的命名。

.....

#### 审稿人 2 意见：

作者的修改很详尽到位，文章同初稿相比有了很大的进步，值得发表。我没有更多的修改意见。谢谢。

非常感谢《心理学报》编辑部和两位评审专家在论文修改过程中给予的帮助和建议！此次论文修改除了参考评审专家的建议，我们还根据学报之前组织的论文写作论坛中的内容，对全文文本进行了挑剔性阅读和修改，以尽可能达到信、达、雅的写作要求。

论文的主要修改如下：

1. 结合审稿意见在研究 1 的讨论中补充了 3-4 岁儿童的线下实验结果，以直接对比线上和线下研究的差异，更全面地总结影响儿童线索偏好表现的因素。

2. 修改了文献综述中有关儿童线索偏好发展的内容，更关注以往研究中 3-8 岁儿童线索偏好的发展介绍，提出了 5-6 岁是儿童社会性发展关键时期这一观点，并据此修改了研究假设。

3. 结合前言的修改和研究 1 中补充的 3-4 岁儿童线下实验结果完善了总讨论。其中，总讨论的第一部分关注了不同实验间 3-8 岁儿童线索偏好结果差异的问题；第二部分总体描述了不同年龄段儿童的线索偏好及其发展特点；第三部分阐述了本研究的理论和实践价值；最后总结了研究不足和展望。

4. 精简和优化了本研究的图表内容和数量。

对于上述内容修改，我们均在正文中以红色文本标记。

---

### 第三轮

#### 审稿人 1 意见：

**意见 1：**既然作者认为 3-4 岁儿童的线下实验结果更为可靠，那为什么还要依赖于线上实验的结果进行分析？如果线上实验可靠性不足，那么研究结果和结论的可靠性也必然会受到影响。研究者应该对实验条件做尽可能好的控制，如有必要作者可考虑补充线下实验的数据，来看研究的结果究竟是怎样的。

**回应：**尊敬的审稿专家，非常感谢您对于本论文的重视和宝贵建议。如您所说，上一轮修改后的论文中确实还存在论文主体和研究结论有一定偏离的问题，我们也对您提出的补充数据的建议进行了讨论和尝试。但很可惜的是，因为本研究对象主要是幼儿，而我们单位所在地的幼儿园和社区都因为一直受疫情影响而采取了严格的防控措施，难以针对研究一的计划样本量（除去已施测的儿童数量后依然接近 180 人）再进行大规模施测。与此同时，在新冠肺炎疫情大流行的背景下，线上的幼儿实验正在成为国内外发展心理学研究者的有效施测方式之一，家长辅助的作用及其价值也逐渐得到了研究者的重视（Rhodes et al., 2020）。在这种情况下，我们最终决定在正文的研究一中保留线上数据的分析，并将相关对研究问题的回答和研究结论限定在线上数据的渠道中。另一方面，为了更好地说明儿童的线索偏好发展的问

题，考虑到研究二中每个年龄段所需的被试数量较少，我们按照您的建议，努力在校园内招募补充了 5-6 岁和 7-8 岁的被试到计划样本量，使之在年龄范围上与研究一对应，保证线下数据也可以反映 3-8 岁儿童的线索偏好发展趋势。

在完成了新的线下数据施测后，我们重新检查了本研究的数据分析方法，发现在分析儿童对群体认知的社会性线索偏好得分时，能够补充将年龄视为连续变量的分析结果，并发现了在排除任务中年龄与社会性线索偏好得分的相关呈边缘显著。与此同时，为了突出研究一的两个实验任务在结果上的异同，我们将先前分开的两个实验合并为了一个研究。此外，研究二在补充了 5-6 岁和 7-8 岁儿童被试之后，我们依据数据特点使用了更直接的列联表卡方检验方法，发现了 3-4 岁和 5-6 岁儿童在颜色×性别×规范条件下的线索偏好差异。总结这两处分析结果的变化，以及不同年龄段儿童在具体实验条件下线索偏好的表现，我们认为，可以从中得出关于儿童的社会性线索偏好随年龄发展而变化的结论。

考虑到儿童社会性线索偏好如何随年龄而发展也是本研究关注的重要问题之一，我们在前言中增加了对这一问题的阐述，补充了相关理论和实证研究，详细说明了以往研究在这一问题上存在的争议，并完善了对本研究假设的论述。通过这种方式，我们希望能够更好地突出研究问题，使前言和讨论更好地呼应，保证文章结构的完整性。

对儿童线下实验结果的补充说明：

在之前的论文版本以及对审稿人的回应中曾详细报告过线上线下实验间的差异，以及线下实验的发现，现对其中重要结果摘录如下：

研究一共在线下招募了 48 名儿童，年龄分布在 3.29-6.92 岁，其中 3-4 岁 35 人，5-6 岁 13 人。对两个年龄段儿童线上和线下实验表现差异的分析发现，线上和线下差异仅在 3-4 岁儿童中显著 ( $t > 3.39$ ,  $p < .002$ )，而在 5-6 岁儿童中并不显著 ( $t < 1.41$ ,  $p > .166$ )。具体而言，当参考正文应用了更为严格的  $p$  值标准 (.006) 时，线下施测的 3-4 岁儿童在研究一两个任务的所有实验条件中都没有表现出基于线索类型的显著偏好 ( $p > .058$ )，而在线上施测中表现出了一定程度的社会性线索偏好 (三元分类任务中有 5 种条件下社会性线索偏好  $p < .002$ ；排除任务中有 3 种条件下存在社会性线索偏好  $p < .001$ )。

根据已有线下实验的结果可以看出，不同于线上实验在 3-4 岁中发现了社会性线索偏好，3-4 岁儿童在线下实验中并没有表现出基于线索的稳定偏好。与此同时，线上和线下的施测方式并没有影响 5-6 岁儿童的线索偏好表现，都发现了他们更偏好使用社会性线索分类。两种施测方式在不同年龄段间的差异印证了之前研究指出的，线上作答仅对 3-4 岁的儿童可能不太准确 (Lapidow et al., 2021)。然而，研究一虽然在线上的测量方法上存在不足，但我们认为这种不足并没有导致非常严重的结果误差。我们做出上述判断主要基于以下两方面原因。

首先，线上与线下实验的结果差异并非本研究关注的重要问题。正如之前审稿过程中提到的，本研究重点想要解决的问题有两方面，分别是 3-8 岁儿童线索偏好情况，以及他们的线索偏好如何在年龄间不同。针对这两个问题，本研究共使用了三种不同的范式进行了检验，相关结果也对研究问题做出了回答，只是因为测量渠道的误差而在一定程度上限制了结果的推广性。在当前论文中，研究一发现 3-4 岁儿童存在社会性线索偏好，但当分析不同年龄段儿童表现出社会性线索偏好的具体实验条件数目可以看出，3-4 岁在社会性线索偏好的稳定性上是低于另外两个较高年龄段儿童的，更容易受到线索情境的影响。此外，当把年龄视为



连续变量，在排除任务中也发现了年龄较大儿童的社会性线索偏好更为显著的趋势（年龄效应呈边缘显著）。结合线下完成的研究二结果可以发现，这种年龄段间线索偏好的发展趋势在不同范式间是相似的，即相比低年龄段，高年龄段儿童对物质性线索的关注降低，而对社会性线索的偏好增强，且年龄差异在 3-4 岁和 5-6 岁之间更为显著。因此，我们认为本研究的两个子研究在年龄如何预测儿童的线索偏好模式上得到了汇聚性的结果，即高年龄段儿童对社会性线索的偏好比低年龄段儿童更强，他们的社会性线索偏好不同实验条件或情境下都更为稳定。

其次，如果以 3-4 岁儿童在研究二中更偏好物质性线索的结果为参照，当前的研究一不论是在线上还是线下的施测结果都没有发现儿童在两项迫选任务中表现出了物质性线索偏好。也就是说，研究一和研究二的结果确实存在差异，且差异的方向不受研究一的线上或线下施测方式的影响，而是研究二的开放性分类任务本身导致儿童在社会分类中更偏好选择物质性线索。针对这一点，我们在讨论中完善了对于实验范式本身如何影响儿童线索偏好选择的解释，认为三选一的开放性分类任务给儿童带来了额外的认知负荷，促使儿童选择他们生活中更熟悉的线索，即性别和经济社会地位，进而导致 3-4 岁儿童在开放性分类任务中表现出了物质性线索偏好。由研究二的方法范式带来的影响同样体现在 5-6 岁和 7-8 岁儿童中，相比于研究一发现的社会性线索偏好，他们在研究二中并没有表现出基于线索类型的显著偏好。而未来的研究可以进一步改进方法技术，在研究中提供更多对应于物质性和社会性线索的线索样例，系统性地检验儿童线索类型偏好的稳定性及其发展。

基于上述原因，我们没有再补充研究一的线下实验被试，而是以现有的实验为核心，在前言中指出了以往研究有关线索偏好发展趋势的争议，并在实验结果基础上针对两个核心问题组织了讨论内容，最后在研究不足中反思了线上和线下施测方式可能给研究结果稳定性带来的影响，指出了采用不同方法检验儿童的线索类型偏好的必要性。本研究所提出的线索类型的框架必然需要更多实证研究的验证和推敲，也需要领域内更多研究者的讨论和建议才能更好地挖掘其价值，这也是我们团队未来前进的方向。

对线上结果有效性的补充说明：

在问卷设计中我们采用了多种方法确保线上测量的有效性。

首先，我们为不同的实验内容设计了家长阅读指导语和为儿童阅读的指导语两部分。在家长部分我们重申了儿童参与反应的重要性，并说明了此测验目的是了解儿童的同伴认知情况，作答没有正确错误之分，家长仅需要阅读指导语让儿童根据问题做出回答。

其次，对于儿童而言，他们需要作答的问题均以图片形式呈现，在每个任务中他们只需要根据指导语要求从 2/3/4 个人物中选出 1 个人物。在一次测验中仅需对 3 类问题作出回答，分别是学习阶段问题、三元分类任务问题和排除任务问题，且每类任务的问题间高度同质。实际上，在线下施测时发现，在实验者说完一遍指导语之后，除个别 3 岁幼儿，几乎所有儿童都能够即时对实验任务做出反应。考虑到实验任务的简单程度，即使是由不熟悉实验方法的家长辅助施测，在已经给予指导语的情况下施测也并不困难。

再次，我们参考线下实验的作答时间设定了有效作答的时间标准（300s），而考虑到每个儿童实际上只需要回答 24 道图片选择问题，且问题间较为同质，这一作答时间限制事实上也较为保守。针对三个年龄段儿童，我们分析了他们作答时间的差异，结果发现不同年龄

段作答时间分布情况相似（如图 R1， $F(2, 212) = 1.661$ ， $p = .192$ ， $\eta^2 = .02$ ），说明儿童在不同年龄段遵循了相似的问答流程，否则更可能出现低年龄段儿童作答时间反而比其他年龄段更短的情况。低年龄段儿童受限于自身对问卷的理解能力，家长在指导时更可能会直接帮助儿童回答从而缩短作答时间，如果事实如此，那么 3-4 岁儿童的作答时间分布应该系统地更为密集，且均值比另外两个年龄段更小，然而当前结果并没有表现出这种趋势。虽然这一分析也并不能直接说明家长完全没有代替儿童作答的情况，但依然从侧面体现了指导语在不同年龄段间的有效性以及施测过程的稳定性。

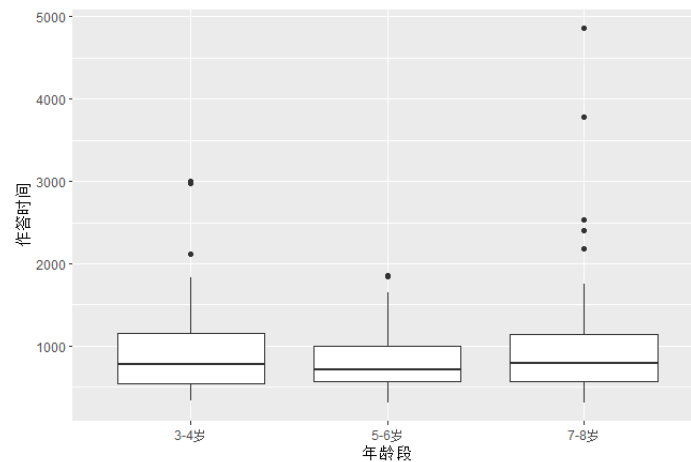


图 R1 不同年龄段儿童作答时间分布

最后，家长参与的儿童实验保证了实验是在儿童更为熟悉和自然的家庭环境中进行施测，没有作为陌生人的研究者在其中的干扰。有研究者指出，这种方式使研究结果在生态性方面更具有优势 (Rhodes et al., 2020)。

总之，结合线上问卷指导语设置情况、问卷难度、答卷筛选标准的选择以及儿童作答环境的情况，我们认为线上作答总体有效，相关实验发现为儿童在群体认知中的线索偏好问题提供了回答，有重要的理论及现实意义。

对于上述说明的内容，正文部分的本次修改已用绿色文字标注，相关文本涉及 前言中 1.3 及 1.4 部分内容，研究一 2.2.1，研究二 3.2.1，讨论 4.1 和 4.2 部分内容。

#### 参考文献：

- Bian, L., Sloane, S., & Baillargeon, R. (2018). Infants expect ingroup support to override fairness when resources are limited. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 2705–2710.
- Lapidow, E., Tandon, T., Goddu, M., & Walker, C. M. (2021). A tale of three platforms: Investigating preschoolers' second-order inferences using in-person, Zoom, and Lookit methodologies. *Frontiers in Psychology*, *12*, 731404.
- Rhodes, M., Rizzo, M. T., Foster-Hanson, E., Moty, K., Leshin, R. A., Wang, M., Benitez, J., & Ocampo, J. D. (2020). Advancing developmental science via unmoderated remote research with children. *Journal of Cognition and Development*, *21*(4): 477-493.

## 第四轮

### 审稿人 1 意见:

作者已对相关问题进行了修改或说明, 没有进一步的问题提出, 建议接受。

### 编委意见:

**意见 1:** 文中的 1.1 和 1.2 采用“社会分类线索”表述, 1.3 采用“群体认知线索”表述。“社会分类线索”和“群体认知线索”是一回事么? 有什么区别与联系? 本文探讨的到底是群体认知线索还是社会分类线索? 请再加以明确。

**回应:** 感谢您的问题与建议。本文探讨的主题是儿童的群体分类线索。使用“社会分类线索”还是“群体认知线索”, 在之前的修改时我们也考虑过但是没有明确说明, 有时候混用, 的确是一个重要问题。在本研究中, 二者指代了同样的线索, 在操作上也是相似的, 都是社会认知线索, 有共同性; 但严格来说两者的涵义有差异。社会分类线索是基于社会分类过程提出的概念, 与社会类别 (social category) 相同, 侧重于对个人或人群形成印象过程中, 对个人或人群身上具有的社会分类线索的简单提取和分类过程, 在此基础上形成了个人印象或群体刻板印象。而群体认知线索是认知群体 (group) 形成、分类与群际关系时使用的线索与信息推断, 所涉及的线索更偏向于是群体构成、对成员而言有实际意义、成员间存在心理依赖性的社会认知线索。一般而言, 社会分类线索包括了个体认知线索与群体认知线索等。

对比两个概念, 群体由具体成员 (个人) 组成, 因此群体认知线索包含对群体类别的异同判断信息, 也包括了对群体具体成员的特征异同与关系的信息。群体形成与分类既可以基于社会分类线索 (如, 性别、种族、社会经济地位、职业等), 也可以来自于社会生活中的不同任务和组织活动。

本研究的核心目的是揭示儿童在群体认知中线索类型重要性的问题, 具体的实验方法上采用了相关的社会分类任务范式。本研究所提出的物质性和社会性线索既是群体认知线索, 也是社会分类线索, 从研究中儿童列举的具体线索中既包括传统的社会分类线索 (如性别、社会经济地位), 也包括了具有实际意义的能够体现成员关系的线索 (如, 兴趣、归属)。结合本研究有关儿童群体认知中线索类型重要性的核心问题, 我们进一步在前言和讨论的有关研究目的和思路中统一了“群体认知线索”或“群体线索”的表述, 并在 1.1 儿童的群体认知线索中陈述了两个概念的关系。

**意见 2:** 根据作者的综述, 1.3 第三自然段中, 对于“两类线索重要性的发展问题”并不构成“矛盾的回答”。以 18 个月的幼儿为被试的研究 (Bian et al., 2018) 结果仅能得出 18 个月幼儿能够意识到社会性线索, 而无法得出“看重社会性线索”的结论, 因为不存在“物质线索”作为对比。

**回应:** 非常感谢您的建议, 我们重新阅读了这里提到的参考文献, 如您所说, 该研究 (Bian et al., 2018) 是以物质性线索的相似性 (相同动物玩具) 区分群体, 研究了幼儿对内群体 vs. 外群体的公平或偏好行为的看法, 并没有呈现物质性线索与社会性线索矛盾的情境。因此, 我们重新组织了 1.3 儿童的群体认知线索偏好及其发展的相关研究部分的写作思路, 首先列举了以往在儿童中比较物质性 vs. 社会性线索的研究内容和发现, 之后总结了研究发现之间的共同点, 同时指出以往研究与当前研究在概念定义上的不同, 以及在参与者年龄范围以及研究方法上的不足之处。这种方式在回应本文研究问题的同时, 也可以为后文的研究假设提供支撑。此外, 也因为研究论据的修改, 我们重新组织了 1.4 当前研究部分的内容。

**意见 3:** 研究一中, 虽然在第一次修改种补充了 G\*power 的一些资料, 但是软件中的 F 检验有很多种, 根据不同的实验设计用法不容, 作者需要详细说明具体用的是哪种。

**回应:** 谢谢您指出该问题, 我们使用的是 F 检验中的 Fixed effects, special, main effects and interactions (ANOVA) 计算了样本量。正文中已做出了对应的修改。

**意见 4:** 研究二中被试数量满足质化研究的样本要求, 是否满足量化研究的样本要求还需要提供证据。

**回应:** 因为我们在最初设计研究二的时候, 更看重质化研究的部分, 因此在招募被试的时候以此为标准收集了数据, 而研究二的量化数据分析也是在修改论文的过程中得以不断完善。针对您的建议, 结合研究目的以及研究二中两部分量化分析的具体结果, 我们对首次社会分类的线索偏好补充了事后检验力的分析, 使用 G\*power 3.1 中  $\chi^2$  检验 Goodness-of-fit tests 进行计算, 发现当前样本在八人和五人分类任务上的检验力近似均为 0.79, 达到了较好的检验力水平。相应内容补充到了 3.2.1 首次分类的线索偏好 中的第一段。

**意见 5:** 研究二的研究设计中, “研究条件”这一被试内变量为什么作者选择(研究条件: 颜色×性别×规范, 经济社会地位×归属×兴趣)这两种组合方式。为什么不是其他线索组合方式, 例如: “颜色×性别×兴趣”等。研究一中表 1 已经得出不同线索组合对不同类型线索偏好有影响, 那么研究二的“研究条件”不同组合方式是否会对结果造成干扰? 这个问题也需要进一步明确。

**回应:** 如您所说, 研究一的发现说明了研究条件这一变量确实会影响儿童线索偏好的结果, 但是为了实现研究目的和实验难度之间的平衡, 我们最终选择了当前的线索组合方式, 具体原因如下。

首先, 我们在设计研究二时也考虑到了研究条件的变量可能会给研究二结果带来的影响, 但由于当时研究二的探索主要以质化分析为主, 因此我们更想从儿童那里获得足够的有关“为什么”选择某一线索分类的信息, 辅以儿童线索类型偏好和线索数量变化的研究结果。

其次, 研究二有效的线索组合方式共有 9 种, 对于每种线索组合, 儿童都要进行 8 次分类(2 组分类对象, 每组分类对象都需要进行 3 次八人分组和 1 次五人分组), 这种实验方式对于儿童来说是不小的负担, 尤其是其中三分之一都是 3-4 岁的幼儿。基于以上原因, 我们最终决定只选择一种线索组合方式向儿童呈现线索。

之所以选择当前的线索组合方式, 主要是基于以往研究对相关线索的研究结果。在选择线索组合的时候, 我们优先考虑的就是希望每组的三个线索对于儿童来说的突显性相当。但是这种控制在不同类型的线索之间显然并不容易, 尤其是在两个研究并行开展的情况下。因此, 我们在设计的时候主要考虑的是这些线索在前人研究中的重要性。在本研究关注的线索中, “性别”是心理学领域最常涉及的主题, 也是对儿童来说非常具有视觉显著性和社会意义的线索(Weisman et al., 2015; Lei et al., 2020), 因此, 对于包含了性别线索的这一组, 我们需要选取另外两个具有代表性和重要性的群体线索。而在社会心理学经典的最简群体范式的研究中, “衣服颜色”是最为常用的区分内外群体的线索之一, 其在儿童研究中的有效性也得到了以往研究的验证(如, Chalik & Rhodes, 2014; Jordan & Dunham, 2020)。因为一组三个线索中需要包含物质性和社会性线索, 最后一个线索需要从社会性线索中选出。相比于归属和兴趣的社会性线索, 规范线索体现了人在认同群体身份基础上遵守了群体的规范性要求, 也与以往研究者对群体社会关系的实验操纵(Bian et al., 2017; Foster-Hanson & Rhodes, 2019; Roberts et al., 2020)有更为直接和紧密的联系。因此, 我们认为颜色和规范在重要性上可以与性别并列, 将其设为其中一个实验条件, 余下的三个群体线索为另一个实验条件。

与上述回应对应, 正文在 3.1 研究方法 部分对实验条件的设定问题进行了补充说明。

虽然在以往研究基础上，我们对群体线索在儿童中的突显性有了大致设想，但从实验结果来看，这种设想并不完善。在较低年龄段的儿童中，每个实验条件下儿童对三种线索的选择比例并非是预期的均分。在性别×颜色×规范组，性别线索在3-4岁儿童中“一枝独秀”，成为儿童群体认知的首要线索偏好；在兴趣×归属×社会经济地位组，3-4岁儿童似乎对归属线索“不屑一顾”，儿童对这一线索的选择概率相对最少。但是，如果我们总结研究二中不同年龄段儿童对每种线索选择偏好的变化趋势，可以大致看出，对于物质性线索，儿童选择颜色线索的概率相对持平，对性别和社会经济地位线索的选择概率相对下跌；对于社会性线索，儿童选择兴趣线索的概率在不同年龄段间大致相同，而对归属和规范线索的选择概率在高年龄段更高。结合研究一中所发现的3-8岁儿童存在社会性线索偏好的结果，我们认为研究二的结果体现了高年龄段儿童的社会性线索偏好更强的趋势，总体支持了研究假设2。

当然，如您所说，当前研究并不能回答，在同时呈现三种线索时，儿童的社会性线索偏好发展是否在不同线索组合条件下稳定的问题。但其实本研究所列举的群体线索也只是无数物质性与社会性线索中的一部分，线索组合在当前的实验设置中是难以穷尽的，这也是本研究在实验方法和设计上的不足之处。因此，我们在4.4 研究不足与展望 的第二段也进行了补充。

---

## 第五轮

### 主编意见：

群体认知是儿童社会性发展的重要内容，不同年龄阶段的儿童在群体社会认知的发展方面有何特点，学界已经进行了较多的研究。作者尝试从物质性线索和社会性线索比较的分类视角开展研究，有一定的理论创新性和应用价值；刻画社会性线索的使用情况和发展趋势也有价值；此外运用多种方法进行检验使得研究结果更加有说服力。建议完善后发表。

**意见 1：** 研究者的自变量主要是物质性和社会性线索，但是结果呈现都是组合效应，研究 1 的实验结果是否应该补充主效应结果趋势？

**回应：** 感谢您的建议。在此基础上，为了更好地在呈现两类线索的效应基础上，展现年龄在其其中的影响，我们采用广义线性混合模型的方法分析了数据，借助 R 4.1.3 软件 (R core team, 2022)，使用 lme4 包 (Bates et al., 2015) 中的 glmer 函数定义了模型，使用 lmerTest 包 (Kuznetsova et al., 2017) 实现的 Satterthwaite 近似获得每个模型生成的参数估计  $p$  值以及固定效应和随机效应结果。具体模型如下：

$$\text{Preference} \sim \text{Age} + \text{Physical} + \text{Societal} + \text{Age:Physical} + \text{Age:Societal} + (1|\text{id})$$

Preference 表示儿童在每个试次中偏好的线索，Age 为作为连续变量的年龄，Physical 为物质性线索变量，Societal 为社会性线索变量，Age:Physical 以及 Age:Societal 分别是年龄与两类线索的交互作用，最后(1|id)为被试的随机效应。

受此模型拟合方法的启发，我们在分析年龄如何在不同实验条件下影响儿童的线索选择这一问题上同样改进了数据分析方法，采用 Logistic 回归，将年龄视为连续变量，从而更好地分析儿童的线索选择随年龄增长而变化的趋势。只在最后为了呼应以往研究对不同年龄段儿童线索类型偏好的研究发现，保留了在具体实验条件下，对不同年龄段间儿童线索偏好表

现的直接比较。对应部分的修改主要体现在研究一的结果部分（2.2.2 儿童线索选择的发展趋势，正文中以蓝色字体表示）。

需要注意的是，新的数据分析方法结果并没有影响研究一的主要发现：①年龄对儿童线索偏好的影响在排除任务中比三元分类任务更为显著；②年龄效应大小在不同的实验条件下有所不同，但对于存在显著年龄效应的实验条件，其都正向预测了儿童对社会性线索的偏好。唯一的差别主要在于，当把年龄视为连续变量，在不同实验条件下进行 Logistic 回归之后，发现了另外两个年龄效应显著的实验条件，分别是三元分类任务中的社会经济地位×归属条件，以及排除任务中的性别×归属条件。在这两种条件下，随年龄增长，儿童的社会性线索偏好有所提升。

此外，广义线性混合模型中的固定效应和随机效应结果并不完全等同于您所提到的物质性线索和社会性线索的主效应，而是同类型的线索样例相互比较的差异显著性。对于两类线索的主效应大小，需要通过模型比较的方式才能得到具体值（此结果已公开上传：[https://osf.io/yq6x5/?view\\_only=bea7823792a54033aaa1ec1f2022c4e4](https://osf.io/yq6x5/?view_only=bea7823792a54033aaa1ec1f2022c4e4)，详见其中的附录文件）。我们最终没有在正文中继续报告主效应结果主要出于以下原因：

从研究问题出发，本研究关注的主要有两方面问题：①儿童的群体认识是更偏好使用物质性线索还是社会性线索；②儿童对于线索类型的偏好如何随年龄发展。其中，对儿童线索类型偏好发展的分析也包含了对儿童在每个年龄段上线索类型偏好的描述。因此，对应于我们的研究问题，年龄效应该是数据分析的重点内容，其效应在社会性线索偏好得分以及具体实验条件下的线索选择上都有所体现。而物质性和社会性线索的主效应对应的是研究选取的物质性（颜色、性别、经济社会地位）和社会性线索样例（规范、兴趣、归属）对儿童线索选择的影响大小。研究中虽然也发现了两类线索存在与年龄的交互效应，但我们更关注的是年龄效应的方向及其稳定性，因此通过具体实验条件下的 Logistic 回归对此进行了检验。相关结果说明不论是在怎样的线索组合条件下，显著的年龄效应的影响方向都是一致的，儿童随年龄增长更多表现出社会性线索偏好。

另一方面，在选择线索样例的时候，我们希望实验涉及的线索是能够充分代表物质性和社会性线索的，预期在两类线索样例的内部存在异质性。在这种情况下，如果我们可以发现儿童存在稳定的线索偏好模式，那就说明我们的研究结果具有一定的推广性，体现了儿童基于两类线索的群体认知可能有本质的不同。而广义线性混合模型发现的物质性和社会性线索及其相关的显著性似乎说明了这一点。在两个实验任务中，线索类型存在主效应和交互效应，与此同时，不论是在怎样的线索组合，不同年龄段儿童也更多地表现出了社会性线索偏好。因此，从这一层面来说，两类线索主效应的大小也并非我们的关注重点，只需要在结果中呈现线索样例确实会对儿童的线索选择产生影响即可。

最后，秉承开放科学的原则，我们公开上传了所有量化研究的数据以及广义线性混合模型分析的代码，并在附录中整合了颜色线索预实验结果、两类线索主效应和交互效应结果、研究一的两个实验任务的所有实验条件下的 Logistic 回归结果（详见：[https://osf.io/yq6x5/?view\\_only=bea7823792a54033aaa1ec1f2022c4e4](https://osf.io/yq6x5/?view_only=bea7823792a54033aaa1ec1f2022c4e4)）。供其他对该主题感兴趣的研究者查阅。

#### 参考文献：

- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Jaeger, B. (2017). r2glmm: Computes R squared for mixed (multilevel) models. *R package version 0.1*, 2, 1-12.
- Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects

Models. *Journal of Statistical Software*, 82(13), 1-26.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

意见 2: 有些文字错误。比如图一中的图 2-1? 图二下面的三元分类任务, 写成了人物?

回应: 感谢指正。我们已重新审阅全文, 以确保不出现错别字和语句方面的问题。