

## 《心理学报》审稿意见与作者回应

题目：3~6 岁幼儿对智能语音助手在不同领域的知识性信任

作者：李喆; 刘浙豫; 毛珂好; 李婉婷; 李婷玉(通讯作者); 李晶

### 第一轮

#### 审稿人 1 意见：

该研究探讨 3-6 岁幼儿对智能语音助手在自然领域和社会领域知识的选择性信任。研究具有一定的理论与现实意义，对于智能时代幼儿获取信息的新途径有一定的指导意义。

**回应：**非常感谢您对本研究的肯定！也非常感谢您的宝贵建议！根据您和另一位审稿专家的建议，我们已逐一修改全文，在此对您的问题进行了回答。

**意见 1：**我的主要问题是关于研究的理论意义，虽然比较智能语音助手与一般成人的选择性信任有一定的现实意义，但是除了选择信任对象的不同，该研究的内容属于经典的选择性信任的问题，研究二探讨的准确性对于选择性信任的影响也已有前人研究，本研究的理论贡献在哪里？作者需要更明确地表明探讨智能语音助手该特殊对象的理论价值。

**回应：**非常感谢您的建议！我们已在文中更明确地探讨本研究的理论价值（见 p2）。

首先，本研究探讨了智能语音助手的准确性是否影响幼儿对其的知识性信任。本文在引言中提到：“由于智能语音助手不具有道德、动机等属性，原则上可能只有一种标准能够适用于智能语音助手，即准确性。有研究表明，幼儿能够基于准确性对电子媒体信息源(如电脑等)进行知识性信任，但就我们所知，目前没有研究直接探讨准确性是否影响幼儿对智能语音助手的知识性信任。”(p3)本研究直接操纵智能语音助手的准确性，结果表明智能语音助手的准确性决定了幼儿对其的知识性信任，明确幼儿对智能语音助手的知识性信任，关键在于智能语音助手的准确性。

其次，本研究探讨幼儿对智能语音助手的知识性信任是否存在领域特殊性。本文在引言中指出：“Girouard-Hallam 和 Danovitch(2022)在研究中将研究问题分为三大维度——稳定的事实类信息、瞬时的事实类信息和个人信息，并发现随着年龄的增长，幼儿对智能语音助手提供的稳定的、瞬时的事实类信息表现出更大的信任，同时在个人信息方面更加偏向于信任人类。但考虑到 Girouard-Hallam 和 Danovitch(2022)的被试选取的局限性，幼儿阶段对智能语音助手在不同领域的知识性信任是否存在差异，仍值得进一步探索。”(p3)本研究考察幼儿对智能语音助手提供的关于自然领域和社会领域(Corriveau et al., 2014; Einav et al., 2013; Clément et al., 2004; 张璟等, 2020)的证言的知识性信任。结果表明“相较于社会领域问题，成人与幼儿在自然领域问题的初始询问中更倾向于选择智能语音助手，这也与我们的假设一致，说明随着年龄的增长和社会经验的增加，幼儿能够判断谁是更好的信息提供者，当涉及到社会领域的问题时，幼儿逐渐认识到具有完整社会属性的人类才能给出更合理的答案。”(p18)

此外，本研究弥补以往研究的不足之处。本文在引言中提到“Girouard-Hallam 和 Danovitch(2022) 在研究中使用的研究范式存在一定缺陷，在其研究一中，智能语音助手和人类对同一问题提供了相同的答案，再由儿童判断信息提供者的回答是否正确。该方法将极大可能影响该研究的效度，一方面，年幼儿童可能会不加选择地认可不同信息提供者的回答(Jaswal et al., 2010); 另一方面，年幼儿童可能会倾向于说‘是’(Okanda & Itakura, 2010), 尽

管 Girouard-Hallam 和 Danovitch(2022)在实验二中将其改为由幼儿选择哪个信息提供者可以给出正确的答案,但这种修改并未从根本上弥补该研究方法的缺陷。”(p4),“同时,该研究的被试年龄跨度较大,存在一定程度的非连续性和不完整性。”(p2),本研究采用知识性信任研究的冲突信息源范式(Koenig et al., 2004; Koenig & Harris, 2005),考察 3~6 岁幼儿对智能语音助手与人类的知识性信任,其中特别关注低龄幼儿(3~4 岁)的证言采纳情况。

综上,“本研究旨在探讨幼儿对智能语音助手的知识性信任的发展状况,并揭示其发展规律。藉由比较智能语音助手与人类作为信息提供者之差异,可为提升人类与科技实体互动之可用性与效用提供参考。探讨幼儿在数字时代的学习生态环境下,如何与智能语音助手互动获取信息和知识,提供家庭和学校使用智能语音助手促进幼儿学习的建议,以及为未来开发智能语音助手提供理论指导。”(p2)

**意见 2:** 第二,本研究发现的领域特殊性主要表现在年幼儿童,年长儿童以及成人在自然与社会领域都会选择信任语音助手,至少在语音助手提供信息准确性高的情况下,作者对此应该做进一步的解释。

**回应:** 非常感谢您的建议!根据您和另一位审稿专家的意见,这部分修改了统计方法,使用 R4.2.2 版中 lme4 软件包中的 glmer 函数和 BOBYQA 优化器建立“广义线性混合模型”(generalized linear mixed model, GLMM)对结果进行统计分析。相应内容已补充在“2.5 统计分析”“2.6 结果与讨论”“3.5 统计分析”“3.6 结果与讨论”中(详见 p8、p14)。

结果表明,幼儿对智能语音助手的知识性信任存在领域特殊性,但在控制智能语音助手和人类的准确性条件下,幼儿与成人在自然和社会领域都会选择准确的信息提供者,我们已对其做进一步解释,补充在“4 讨论”中(详见 p17)。

“实验一结果表明,相较于社会领域问题,成人与幼儿在自然领域问题的初始询问中更倾向于选择智能语音助手,5 岁、6 岁幼儿与成人在自然领域对智能语音助手的平均询问意愿、平均选择信任和平均最终判断均高于社会领域。这与我们的假设一致,说明随着年龄的增长和社会经验的增加,幼儿能够判断谁是更好的信息提供者,当涉及到社会领域的问题时,幼儿逐渐认识到具有完整社会属性的人类才能给出更合理的答案,或者说幼儿认识到人类才是社会领域中的专家。同时,4 岁幼儿对智能语音助手在自然领域的平均询问意愿和平均选择信任高于社会领域,但在最终判断中,4 岁幼儿表现出了对智能语音助手在自然领域的信任偏好,研究者推测出现这种左右摇摆的情况可能是因为社会领域的任务比自然领域的任务更为复杂,4 岁幼儿的认知发展还不足以让他们准确、快速判断获得的信息。这个发现也表明 4 岁幼儿已经能对获得的信息进行复杂的分析判断,认识到不同的信息提供者有不同的领域专业性,并以此为依据选择信任领域专业性更好的信息提供者。

而在操纵智能语音助手和人类准确性的条件下,成人与幼儿的选择均不存在领域特殊性,在自然领域和社会领域中,成人与幼儿更倾向于选择信任准确的信息提供者。这从另一个角度说明准确性可能是幼儿对其知识性信任的先决条件,准确性是幼儿更重视的好的信息提供者的特征。”

**意见 3:** 第三,研究二为什么要把年龄扩展到 3、4、5 岁,而不是研究一的 4、5、6 岁,作者需要做解释。

**回应:** 非常感谢您的建议!我们已经对此做出更明确的解释。相关内容已补充在文中 p11“以上结果说明,……鉴于实验一的结果显示,4 岁幼儿已具备与年长幼儿和成人类似的理性的知识性信任决策能力,而以往研究发现 4 岁以下幼儿更可能存在轻信倾向(Jaswal et al., 2010; Jaswal et al., 2014; Tong et al., 2020)。为了考察幼儿对智能语音助手知识性信任的发展规律,有必要从更早(3 岁)开始连续考察幼儿对智能语音助手知识性信任的发展情况;实验二将被

试年龄向下扩展到 3 岁，以探究更年幼的幼儿对智能语音助手的知识性信任。”

**意见 4:** 第四，文中有很多语句不通顺的地方，比如，即更容易不加选择的相信他人他们被告知的内容，倾向于从信任之前准确的电脑而非不准确的电脑呈现的信息，最后结论的多余句号等等，请作者仔细阅读全文，避免低级错误。

**回应:** 非常感谢您的建议！十分惭愧！我们已仔细阅读全文，对文中有错别字的地方、格式有误的地方以及语句不通顺的地方进行修改，例如，将您在此列出的“即更容易不加选择的相信他人他们被告知的内容”，改为“即更容易不加选择地相信他人提供的信息”；(p2) “倾向于从信任之前准确的电脑而非不准确的电脑呈现的信息”改为“倾向于信任以前准确的电脑提供的信息，而不相信错误的电脑。”；(p3) “最后结论的多余句号”已经去除。(p19)

**审稿人 2 意见:**

《3~6 岁幼儿对智能语音助手在不同领域的知识性信任》一文的选题，具有一定的创新性。不过，论文还存在较多问题，质量亟待提升。

**回应:** 非常感谢您的宝贵建议！我们已根据您的建议逐一修改全文，在此先对您的问题进行回答。

**意见 1: 实验二的设计令人存在诸多疑惑。**

**疑惑 1:** 该实验设计跟作者在研究程序中的介绍不一致，因此到底以实验设计为准还是以研究程序中的介绍为准，令人疑惑。

以下为论文中关于实验二研究程序的介绍：

“主试告诉幼儿自己会拿出几张图片，‘小爱’和‘圆圆’会告诉他们这是什么。幼儿被随机分配到‘圆圆’准确或‘小爱’准确的条件下，在‘圆圆’准确条件下，主试告诉幼儿每次‘圆圆’都会提供‘对的’的回答，而‘小爱’则会提供‘错的’回答。反之，对‘小爱’准确条件下的幼儿来说，‘小爱’是可靠的信息来源。”由此可见，作者似乎同时操控了智能语音助手和人类信息提供者的准确性，而且将这两个信息提供者特意设定为“此对彼错”的状态。

以下为论文中实验二的设计：

“实验采用 3（年龄：3 岁/4 岁/5 岁） $\times$ 2（“小爱”的准确性：准确 VS 不准确） $\times$ 2（信息提供者：人类 VS 智能语音系统） $\times$ 2 问题的领域：自然 VS 社会）四因素混合实验设计，年龄为组间自变量，信息提供者和问题领域为组内自变量，因变量为被试对信息提供者的询问选择、对信息提供者的证言选择以及被试最终判断（是 VS 否）与智能语音助手‘小爱’一致。”

从这个实验设计来看，作者只控制了智能语音助手“小爱”的准确性，并没有提及人类信息提供者的准确性。既然如此，请问作者是如何完成研究程序中“此对彼错”的操作呢？

**回应:** 非常感谢您为我们指出这一问题！实验二实际上是同时控制不同信息源类型(智能语音助手 VS 人类)的准确性，很抱歉我们在上一稿中没有阐释清楚，我们已将其修改为“实验采用 4(年龄：3 岁/4 岁/5 岁/成人) $\times$ 2(准确性条件：“小爱不准确，圆圆准确”VS“小爱准确，圆圆不准确”) $\times$ 2(问题的领域：自然 VS 社会)三因素混合实验设计，其中准确性条件从 2 准确性(准确 VS 不准确)  $\times$ 2 信息提供者(人类 VS 智能语音系统)，四种可能情况中选择出来，考虑到本实验操纵的是两种信息源准确性的差异，人类与智能语音助手都准确或都不准确的情况被排除。”(p13)

**疑惑 2:** 实验设计中总共四个自变量，只介绍了“年龄为组间自变量，信息提供者和问

题领域为组内自变量”，并未介绍另一个自变量“‘小爱’的准确性：准确 VS 不准确”是被试内还是被试间自变量。无论是被试内还是被试间自变量，似乎都跟后续的统计结果难以对应上。

回应：答：非常感谢您为我们指出这一问题，很抱歉我们之前的表述有误，已经更正。(p13)同时，根据您的建议，我们修改了本研究的统计方法。本研究使用 R4.2.2 版中 lme4 软件包中的 glmer 函数和 BOBYQA 优化器建立“广义线性混合模型”(generalized linear mixed model, GLMM)对结果进行统计分析。相应内容已补充在“2.5 统计分析”“2.6 结果与讨论”“3.5 统计分析”“3.6 结果与讨论”中(详见 p8、p14)。

实验二的统计修改部分如下所示：

“将幼儿对不同准确性(准确 VS 不准确)的信息提供者在不同领域(自然 VS 社会)的知识性信任(询问意愿)作为因变量, GLMM 全模型包含以下固定效应: 准确性条件、年龄、领域以及三者的两两交互作用, 随机变量包括幼儿 ID 和试次, 需要特别说明的是, 我们曾尝试将准确性条件、年龄、领域的三重交互纳入模型, 但发现模型无法拟合, 因此并未将其纳入最终模型。通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 0.05, df = 1, p = 0.815$ ), 年龄主效应不显著( $\chi^2 = 1.55, df = 3, p = 0.672$ ), 准确性条件主效应显著( $\chi^2 = 97.57, df = 1, p < 0.001$ ), 准确性条件与年龄交互作用显著( $\chi^2 = 25.66, df = 3, p < 0.001$ ), 准确性条件与领域( $\chi^2 = 0.11, df = 1, p = 0.738$ )、年龄与领域( $\chi^2 = 1.81, df = 3, p = 0.613$ )均不存在交互作用。准确性条件与年龄的交互作用如图 9 所示。对准确性条件进行事后对比发现, 幼儿更愿意选择询问准确的信息提供者( $t = 28.62, p < 0.001$ ), 不考虑是人还是智能语音助手。

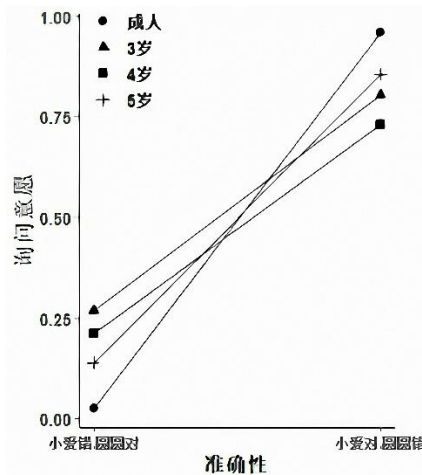


图 9 准确性与年龄的交互作用

进一步对准确性条件和年龄的交互作用进行简单效应分析发现, 成人与三个年龄段幼儿均更愿意选择询问准确信息提供者(成人:  $p < 0.001$ ; 3 岁:  $p < 0.001$ ; 4 岁:  $p < 0.001$ ; 5 岁:  $p < 0.001$ )。在小爱不准确, 圆圆准确条件下, 相较于三个年龄段的幼儿, 成人更愿意选择询问准确信息提供者(成人-3 岁:  $p = 0.010$ ; 成人-4 岁:  $p = 0.002 < 0.001$ ; 成人-5 岁:  $p = 0.033 < 0.05$ ), 各年龄组幼儿间的选择无显著差异(3 岁-4 岁:  $p = 0.999$ ; 3 岁-5 岁:  $p = 0.725$ ; 4 岁-5 岁:  $p = 0.606$ )。在小爱准确, 圆圆不准确条件下, 相较于 4 岁幼儿, 成人更愿意选择询问准确信息提供者( $p < 0.001$ ); 各年龄组幼儿间的选择无显著差异(3 岁-4 岁:  $p = 0.406$ ; 3 岁-5 岁:  $p = 0.999$ ; 4 岁-5 岁:  $p = 0.265$ )。

以上结果说明, 在操纵准确性的条件下, 成人与幼儿均更愿意选择询问准确信息提供者。相较于幼儿, 成人更愿意选择询问准确信息提供者。”(p14)

“将幼儿对不同准确性(准确 VS 不准确)的信息提供者在不同领域(自然 VS 社会)的知识性信任(外显信任)作为因变量, GLMM 全模型包含以下固定效应: 准确性条件、年龄、领域以及三者的两两交互作用, 随机变量包括幼儿 ID 和试次, 通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 0.26, df = 1, p = 0.612$ ), 年龄主效应显著( $\chi^2 = 8.10, df = 3, p = 0.044 < 0.05$ ), 准确性条件主效应显著( $\chi^2 = 103.29, df = 1, p < 0.001$ ), 准确性条件与年龄( $\chi^2 = 1.29, df = 3, p = 0.731$ )、准确性条件与领域( $\chi^2 = 3.13, df = 1, p = 0.077$ )、年龄与领域( $\chi^2 = 1.16, df = 3, p = 0.762$ )均不存在交互作用。对年龄进一步进行事后比较发现, 相较于 3 岁幼儿, 4 岁、5 岁幼儿和成人均更倾向于选择信任准确的智能语音助手(4 岁-3 岁:  $t = 2.85, p = 0.004 < 0.01$ ; 5 岁-3 岁:  $t = 3.95, p < 0.001$ ; 成人-3 岁:  $t = 2.02, p = 0.043 < 0.05$ )。对准确性条件进行事后比较发现, 幼儿更倾向于选择信任准确的信息提供者( $t = 41.27, p < 0.001$ )。

以上结果说明, 在操纵准确性的条件下, 成人与幼儿均更愿意选择信任准确信息提供者。相较于 3 岁幼儿, 4 岁、5 岁幼儿和成人均更倾向于选择信任准确的智能语音助手。”(p16)

“将幼儿对不同准确性(准确 VS 不准确)的信息提供者在不同领域(自然 VS 社会)的知识性信任(最终判断)作为因变量, GLMM 全模型包含以下固定效应: 准确性条件、年龄、领域以及三者的两两交互作用, 随机变量包括幼儿 ID 和试次, 通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 0.19, df = 1, p = 0.662$ ), 年龄主效应不显著( $\chi^2 = 3.96, df = 3, p = 0.265$ ), 准确性条件主效应显著( $\chi^2 = 88.39, df = 1, p < 0.001$ ), 准确性条件与年龄( $\chi^2 = 2.81, df = 3, p = 0.422$ )、准确性条件与领域( $\chi^2 = 3.79, df = 1, p = 0.051$ )、年龄与领域( $\chi^2 = 0.60, df = 3, p = 0.897$ )均不存在交互作用。对准确性条件进行事后比较发现, 幼儿更倾向于最终信任准确的信息提供者( $t = 42.79, p < 0.001$ )。

以上结果说明, 在操纵准确性的条件下, 成人与幼儿均更愿意选择信任准确信息提供者。”(p16)

疑惑 3: 若推测自变量“‘小爱’的准确性: 准确 VS 不准确”, 实际上为“信息的准确性: 准确 VS 不准确”之误, 在其他自变量不作任何改动的情况下, 就目前实验二的设计来看, 那也跟研究程序和统计数据对应不上, 这是为何?

回应: 非常感谢您指出的这一问题! 自变量“‘小爱’的准确性: 准确 VS 不准确”, 实际上为“‘小爱’与‘圆圆’的准确性: 小爱不准确, 圆圆准确 VS 小爱准确, 圆圆不准确”, 根据您的建议, 我们已修改实验设计的阐述和统计分析方式, 希望能解答您的疑惑。

疑惑 4: 实验二的设计中提到的四个自变量, 并没有“性别”, 在三次的重复测量方差分析中也没有提及性别主效应如何, 但是在“3.6.3 幼儿对信息提供者的最终采信”这个部分, 突然提到“准确性和性别的交互效应显著”。若“性别”也是自变量, 为何实验设计中不提? 引言中也并未提及关注性别影响的有关文献, 这是为何?

回应: 非常感谢您提出的这一问题! 诚如您所说, 性别并不是我们关键和操纵的自变量, 我们在实验一和实验二选取的各年龄段被试均男女各半。在之前的重复测量方差分析中, 我们将性别放入其中分析其实是想排除这一无关变量的影响, 但却意外地发现在实验二的最终采信部分, 准确性与性别的交互作用显著, 于是我们在文中进行了报告。根据您的和另一位审稿专家的意见, 我们参考相关文献, 将实验一、实验二的统计方式改为利用 R4.2.2 建立 GLMM 模型, 在我们建立模型的过程中, 也曾尝试将性别作为随机变量纳入模型, 但发现并未改善模型的拟合度, 同时部分模型出现无法拟合的情况, 考虑到性别并不是我们实验的自变量, 因此并未将其纳入最终模型中。

所以，作者到底是如何做了实验二，实在令人不解。如此，实验二的研究结果和结论，也暂不好研判是否可靠和可信。

回应：非常感谢您对我们实验二提出的相关问题！实验一的结果表明，“相较于社会领域问题，成人与幼儿在自然领域问题的初始询问中更倾向于选择智能语音助手。同时，随着年龄的增长，幼儿越来越倾向于选择信任智能语音助手，相比于4、5岁组幼儿，6岁组幼儿更愿意选择信任智能语音助手；相较于幼儿，成人更愿意选择信任智能语音助手。这与预期一致，说明随着年龄的增长和社会经验的增加，幼儿能够判断谁是更好的信息提供者。此外，结果表明4岁幼儿已经能对获得的信息进行复杂的分析判断” (p11)“考虑到已有的基于人类信息提供者的研究发现4岁以下的幼儿更有可能表现出轻信偏见(bias to believe)，即更容易不加选择地相信他人提供的信息(甚至在有明显提示证言为错误的情况下)(Jaswal et al., 2010; Jaswal et al., 2014; Tong et al., 2020)。” (p2)我们认为有必要从更年幼的幼儿(3岁幼儿)开始连续考察幼儿期儿童对智能语音助手的知识性信任的发展情况。因此，在控制智能语音助手和人类的准确性条件下，我们将年龄扩展到3、4、5岁，以更完整地探究幼儿对智能语音助手的知识性信任，并重点关注3~4岁的低龄幼儿的证言采纳情况。相关内容已补充在文中 p11。

我们已根据您的建议，并查阅相关文献，对我们阐述不当的地方逐一修改，同时为使实验过程更加清晰，我们修改了实验过程部分的阐述，并增加了实验流程图，见文 p8、p12，也希望上述回答能解答您的疑惑。再次感谢您提出的相关问题！

### “3.3 研究过程

与实验一基本相同，仅在热身阶段后加入熟悉阶段，并在熟悉阶段中操纵信息提供者的准确性历史(见图7)。

#### 热身阶段

与实验一相同。

#### 熟悉阶段

主试告诉幼儿自己会拿出几张图片，“小爱”和“圆圆”会告诉他们这是什么。幼儿被随机分配到“小爱不准确，圆圆准确”或“小爱准确，圆圆不准确”的条件下，在“小爱不准确，圆圆准确”下，主试告诉幼儿每次“圆圆”都会提供“对的”的回答，而“小爱”则会提供“错的”回答。反之，对“小爱准确，圆圆不准确”条件下的幼儿来说，“小爱”是可靠的信息来源。

接下来，主试连续出示三张熟悉物品图片(见表3)，并让两个信息提供者告知幼儿该物品的名称，准确的信息提供者总是回答正确，不准确的信息提供者总是回答错误，例如主试出示“裤子”的图片，准确的信息提供者指出它是“裤子”，而不准确的信息提供者则回答“鞋子”。实验中图片出现的顺序、信息提供者的位置、信息源的准确性、证言的提供顺序等均经过事先的平衡。

测试阶段和去处理阶段 与实验一相同。”(p12)

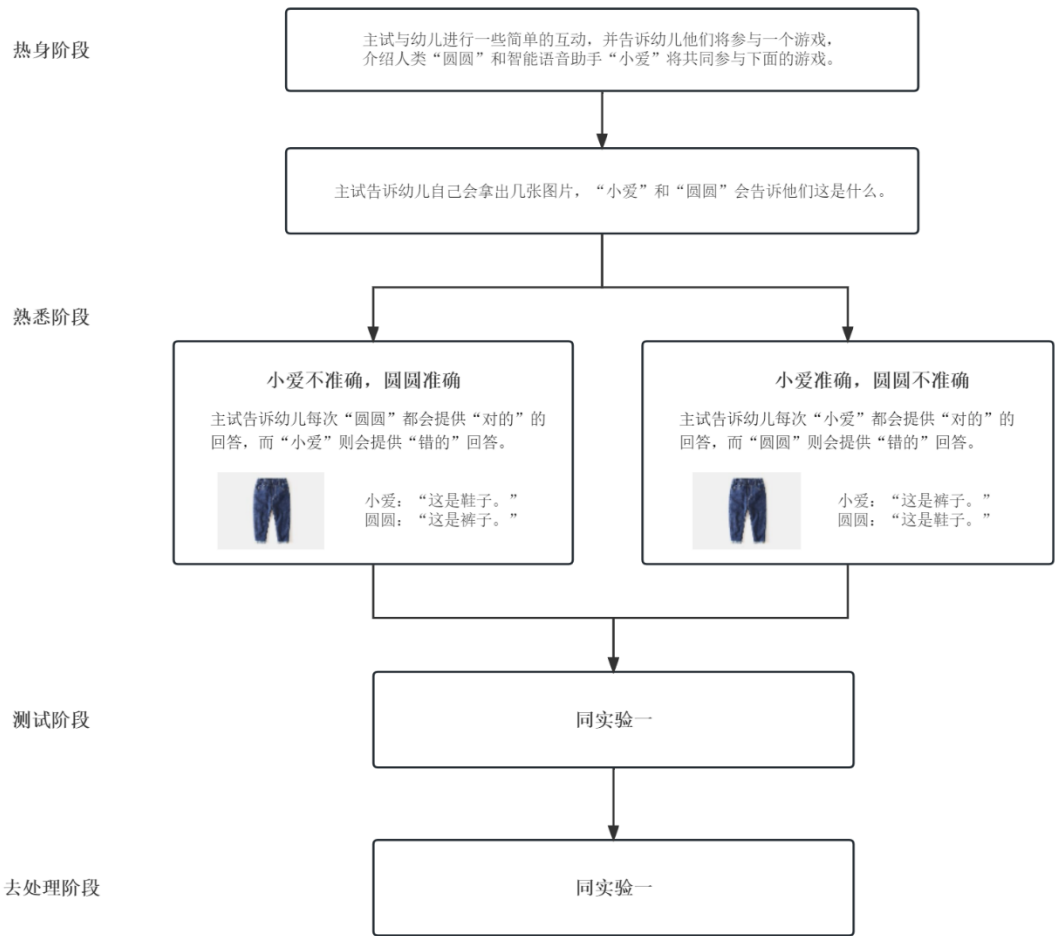


图 7 实验二研究过程流程图

意见 2: 两个实验数据统计的方式令人疑惑。

实验一和实验二的结果部分，为何均对于同一批数据，既进行非参数检验（卡方检验）又进行参数检验（重复测量方差分析），请问作者有何特别的考虑？若用于重复测量方差分析的数据，并非论文图中的同一批数据，请问重复测量方差分析的数据又在哪里？

回应: 非常感谢您提出的问题! 在之前的分析中，我们之所以进行与随机概率比的卡方检验，是因为想说明幼儿的选择并不是随机选的，根据您的建议，我们已经修改统计方式，将重复测量方差分析改为建立 GLMM 模型，以进行更为完整、准确的统计分析。同时，我们将卡方检验改为对幼儿和成人的平均选择进行统计分析，并将其绘成 error bars 图，以更直观地说明在幼儿与成人的选择。(详见 p9、p14)

例如实验一中:

“对幼儿及成人在自然领域和社会领域对智能语音助手的平均询问意愿进行分析，结果如图 4 所示。

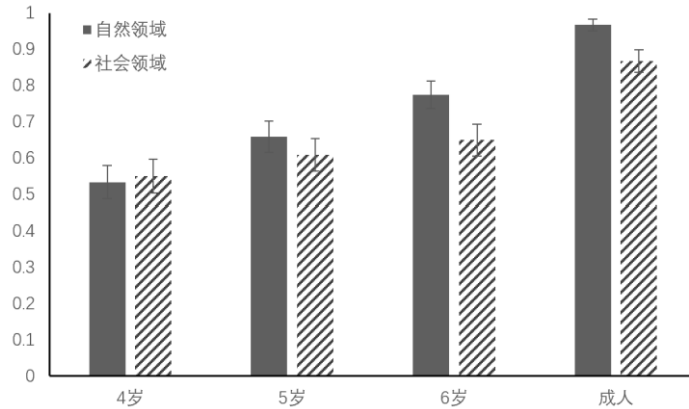


图4 幼儿与成人在不同领域对智能语音助手的平均询问意愿

将幼儿对智能语音助手在不同领域(自然 VS 社会)的知识性信任(询问意愿)作为因变量, GLMM 全模型包含以下固定效应(fixed effect): 年龄、领域、年龄和领域的交互作用, 随机变量(random factor)包括幼儿 ID 和试次。通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应显著( $\chi^2 = 3.85, df = 1, p = 0.050$ ), 年龄主效应显著( $\chi^2 = 60.12, df = 3, p < 0.001$ ), 领域和年龄的交互作用不显著( $\chi^2 = 3.85, df = 3, p = 0.084$ )。对领域进行事后对比发现, 幼儿在社会领域选择询问智能语音助手的意愿显著低于自然领域( $t = 2.19, p = 0.029 < 0.05$ )。对年龄进一步进行事后对比发现, 相较于幼儿, 成人更愿意选择询问智能语音助手(成人-4岁:  $t = 9.39, p < 0.001$ ; 成人-5岁:  $t = 7.10, p < 0.001$ ; 成人-6岁:  $t = 5.11, p < 0.001$ ), 年龄越大的幼儿越倾向于选择询问智能语音助手(5岁-4岁:  $t = 2.30, p = 0.022$ ; 6岁-4岁:  $t = 4.28, p < 0.001$ ; 6岁-5岁:  $t = 1.98, p = 0.048 < 0.05$ )。

以上结果说明, 成人与幼儿对智能语音助手的初始信任偏好受到不同问题领域的影响, 相较于社会领域问题, 成人与幼儿在自然领域问题的初始询问中均更倾向选择智能语音助手。同时, 相较于幼儿, 成人在初始询问中更愿意选择智能语音助手, 年龄越大的幼儿越倾向于选择询问智能语音助手。”(p9)

“对幼儿及成人在自然领域和社会领域对智能语音助手的平均外显信任进行分析, 结果如图5所示。

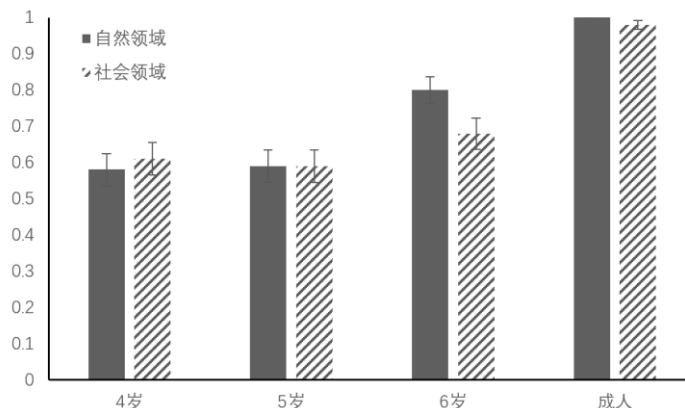


图5 幼儿与成人在不同领域对智能语音助手的平均外显信任

将幼儿对智能语音助手在不同领域(自然 VS 社会)的知识性信任(外显信任)作为因变量, GLMM 全模型包含以下固定效应: 年龄、领域、年龄和领域的交互作用, 随机变量包括幼



儿 ID 和试次, 通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 0.45, df = 1, p = 0.502$ ), 年龄主效应显著( $\chi^2 = 32.37, df = 3, p < 0.001$ ), 领域和年龄的交互作用不显著( $\chi^2 = 2.68, df = 3, p = 0.443$ )。对年龄进一步进行事后对比发现, 相较于幼儿, 成人更愿意选择信任智能语音助手(成人-4岁:  $t = 10.48, p < 0.001$ ; 成人-5岁:  $t = 10.59, p < 0.001$ ; 成人-6岁:  $t = 6.62, p < 0.001$ ); 相较于4、5岁幼儿, 6岁幼儿更愿意选择信任智能语音助手(6岁-4岁:  $t = 3.86, p < 0.001$ ; 6岁-5岁:  $t = 3.97, p < 0.001$ )。

以上结果说明, 相较于幼儿, 成人更愿意选择信任智能语音助手, 相较于4岁和5岁幼儿, 6岁幼儿更愿意选择信任智能语音助手。”(p10)

“对幼儿及成人在自然领域和社会领域对智能语音助手的平均最终判断进行分析, 结果如图6所示。

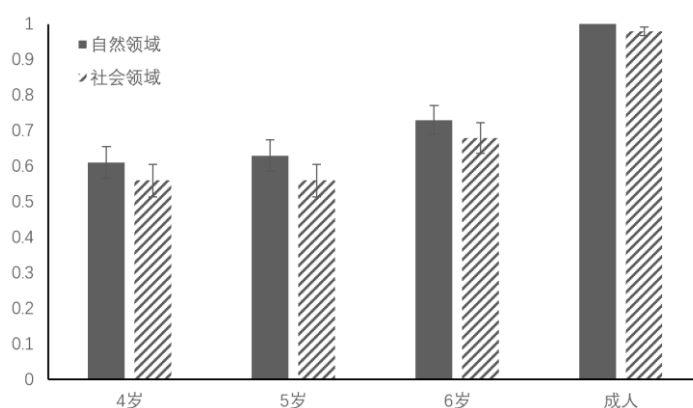


图6 幼儿与成人在不同领域对智能语音助手的平均最终判断

将幼儿对智能语音助手在不同领域(自然 VS 社会)的知识性信任(最终判断)作为因变量, GLMM 全模型包含以下固定效应: 年龄、领域、年龄和领域的交互作用, 随机变量包括幼儿 ID 和试次, 通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 1.81, df = 1, p = 0.178$ ), 年龄主效应显著( $\chi^2 = 33.49, df = 3, p < 0.001$ ), 领域和年龄的交互作用不显著( $\chi^2 = 0.20, df = 3, p = 0.978$ )。对年龄进一步进行事后对比发现, 相较于幼儿, 成人更愿意最终信任智能语音助手(成人-4岁:  $t = 10.67, p < 0.001$ ; 成人-5岁:  $t = 10.46, p < 0.001$ ; 成人-6岁:  $t = 7.41, p < 0.001$ ); 相较于4、5岁幼儿, 6岁幼儿更愿意最终信任智能语音助手(6岁-4岁:  $t = 3.27, p = 0.001$ ; 6岁-5岁:  $t = 3.05, p = 0.002 < 0.01$ )。

以上结果说明, 相较于幼儿, 成人更愿意最终接受智能语音助手的证言; 同时, 6岁幼儿的最终判断与外显信任保持一致, 相比于4、5岁组幼儿, 6岁组幼儿更愿意最终接受智能语音助手的证言。”(p11)

例如实验二中:

“在不同准确性条件下, 对幼儿及成人在自然领域和社会领域对智能语音助手的平均询问意愿进行分析, 结果如图8所示。当智能语音助手失去其准确性时, 三个年龄段的幼儿及成人对智能语音助手在自然领域和社会领域的平均询问意愿均下降。

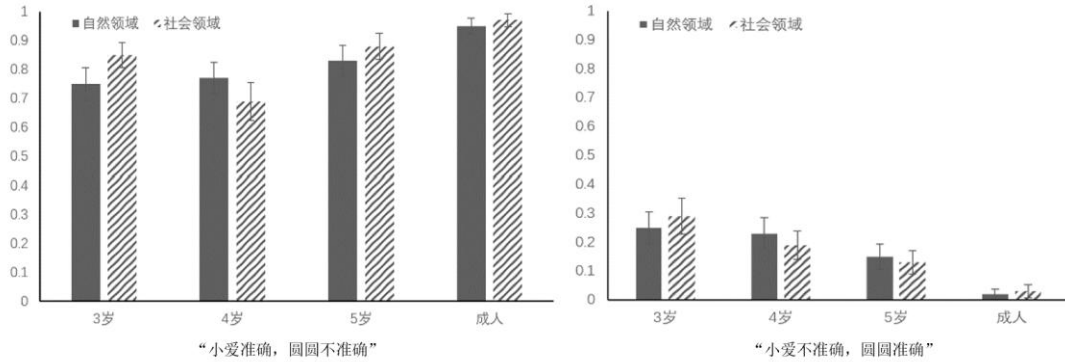


图8 不同准确性条件下, 幼儿与成人在不同领域对智能语音助手的平均询问意愿

将幼儿对不同准确性(准确 VS 不准确)的信息提供者在不同领域(自然 VS 社会)的知识性信任(询问意愿)作为因变量, GLMM 全模型包含以下固定效应: 准确性条件、年龄、领域以及三者的两两交互作用, 随机变量包括幼儿 ID 和试次, 需要特别说明的是, 我们曾尝试将准确性条件、年龄、领域的三重交互纳入模型, 但发现模型无法拟合, 因此并未将其纳入最终模型。通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 0.05, df = 1, p = 0.815$ ), 年龄主效应不显著( $\chi^2 = 1.55, df = 3, p = 0.672$ ), 准确性条件主效应显著( $\chi^2 = 97.57, df = 1, p < 0.001$ ), 准确性条件与年龄交互作用显著( $\chi^2 = 25.66, df = 3, p < 0.001$ ), 准确性条件与领域( $\chi^2 = 0.11, df = 1, p = 0.738$ )、年龄与领域( $\chi^2 = 1.81, df = 3, p = 0.613$ )均不存在交互作用。准确性条件与年龄的交互作用如图 9 所示。对准确性条件进行事后对比发现, 幼儿更愿意选择询问准确的信息提供者( $t = 28.62, p < 0.001$ ), 不考虑是人还是智能语音助手。

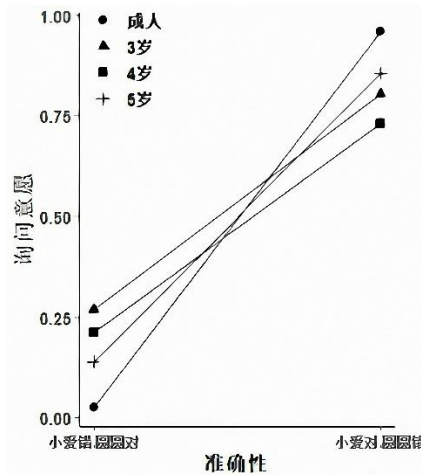


图9 准确性条件与年龄的交互作用

进一步对准确性条件和年龄的交互作用进行简单效应分析发现, 成人与三个年龄段幼儿均更愿意选择询问准确信息提供者(成人:  $p < 0.001$ ; 3岁:  $p < 0.001$ ; 4岁:  $p < 0.001$ ; 5岁:  $p < 0.001$ )。在小爱不准确, 圆圆准确条件下, 相较于三个年龄段的幼儿, 成人更愿意选择询问准确信息提供者(成人-3岁:  $p = 0.010$ ; 成人-4岁:  $p = 0.002 < 0.001$ ; 成人-5岁:  $p = 0.033 < 0.05$ ), 各年龄组幼儿间的选择无显著差异(3岁-4岁:  $p = 0.999$ ; 3岁-5岁:  $p = 0.725$ ; 4岁-5岁:  $p = 0.606$ )。在小爱准确, 圆圆不准确条件下, 相较于4岁幼儿, 成人更愿意选择询问准确信

息提供者( $p < 0.001$ ); 各年龄组幼儿间的选择无显著差异(3岁-4岁:  $p = 0.406$ ; 3岁-5岁:  $p = 0.999$ ; 4岁-5岁:  $p = 0.265$ )。

以上结果说明,在操纵准确性的条件下,成人与幼儿均更愿意选择询问准确信息提供者。相较于幼儿,成人更愿意选择询问准确信息提供者。”(p14)

“在不同准确性条件下,对幼儿及成人在自然领域和社会领域对智能语音助手的平均外显信任进行分析,结果如图 10 所示。当智能语音助手失去其准确性时,三个年龄段的幼儿及成人对智能语音助手在自然领域和社会领域的平均外显信任均下降。

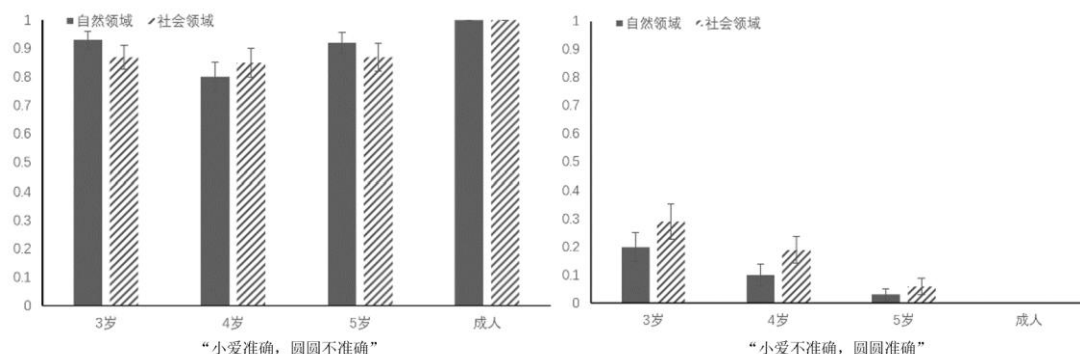


图 10 不同准确性条件下,幼儿与成人在不同领域对智能语音助手的平均外显信任

将幼儿对不同准确性(准确 VS 不准确)的信息提供者在不同领域(自然 VS 社会)的知识性信任(外显信任)作为因变量, GLMM 全模型包含以下固定效应:准确性条件、年龄、领域以及三者的两两交互作用,随机变量包括幼儿 ID 和试次,通过主成分分析法逐一缩减随机斜率与随机截距,最终模型的结果表明领域主效应不显著( $\chi^2 = 0.26, df = 1, p = 0.612$ ),年龄主效应显著( $\chi^2 = 8.10, df = 3, p = 0.044 < 0.05$ ),准确性条件主效应显著( $\chi^2 = 103.29, df = 1, p < 0.001$ ),准确性条件与年龄( $\chi^2 = 1.29, df = 3, p = 0.731$ )、准确性条件与领域( $\chi^2 = 3.13, df = 1, p = 0.077$ )、年龄与领域( $\chi^2 = 1.16, df = 3, p = 0.762$ )均不存在交互作用。对年龄进一步进行事后比较发现,相较于 3 岁幼儿,4 岁、5 岁幼儿和成人均更倾向于选择信任准确的智能语音助手(4岁-3岁:  $t = 2.85, p = 0.004 < 0.01$ ; 5岁-3岁:  $t = 3.95, p < 0.001$ ; 成人-3岁:  $t = 2.02, p = 0.043 < 0.05$ )。对准确性条件进行事后比较发现,幼儿更倾向于选择信任准确的信息提供者( $t = 41.27, p < 0.001$ )。

以上结果说明,在操纵准确性的条件下,成人与幼儿均更愿意选择信任准确信息提供者。相较于 3 岁幼儿,4 岁、5 岁幼儿和成人均更倾向于选择信任准确的智能语音助手。”(p16)

“在不同准确性条件下,对幼儿及成人在自然领域和社会领域对智能语音助手的平均最终判断进行分析,结果如图 11 所示。当智能语音助手失去其准确性时,三个年龄段的幼儿及成人对智能语音助手在自然领域和社会领域的平均最终判断均下降。

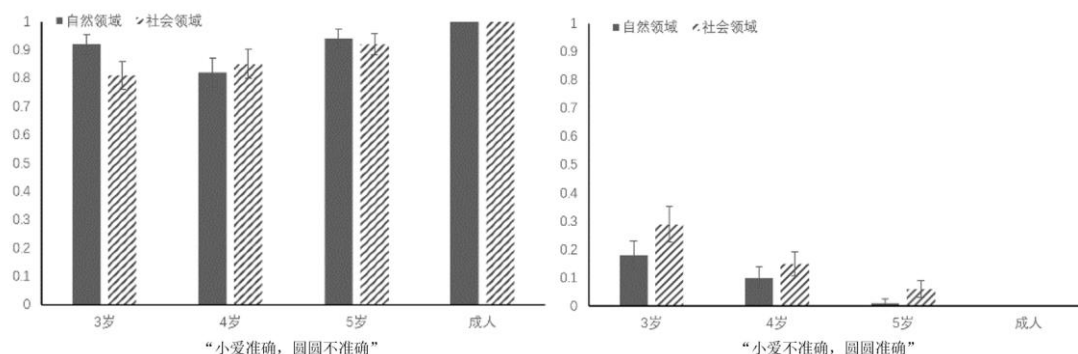


图 11 不同准确性条件下,幼儿与成人在不同领域对智能语音助手的平均最终判断

将幼儿对不同准确性(准确 VS 不准确)的信息提供者在不同领域(自然 VS 社会)的知识性信任(最终判断)作为因变量, GLMM 全模型包含以下固定效应: 准确性条件、年龄、领域以及三者的两两交互作用, 随机变量包括幼儿 ID 和试次, 通过主成分分析法逐一缩减随机斜率与随机截距, 最终模型的结果表明领域主效应不显著( $\chi^2 = 0.19, df = 1, p = 0.662$ ), 年龄主效应不显著( $\chi^2 = 3.96, df = 3, p = 0.265$ ), 准确性条件主效应显著( $\chi^2 = 88.39, df = 1, p < 0.001$ ), 准确性条件与年龄( $\chi^2 = 2.81, df = 3, p = 0.422$ )、准确性条件与领域( $\chi^2 = 3.79, df = 1, p = 0.051$ )、年龄与领域( $\chi^2 = 0.60, df = 3, p = 0.897$ )均不存在交互作用。对准确性条件进行事后比较发现, 幼儿更倾向于最终信任准确的信息提供者( $t = 42.79, p < 0.001$ )。

以上结果说明, 在操纵准确性的条件下, 成人与幼儿均更愿意选择信任准确信息提供者。”(p16)

**意见 3:** 研究被试群体和论文内容之间不匹配。

实验一和实验二中均有成人被试, 但是论文题目、摘要、两个实验的设计和结论中均没有成人被试及其研究结果的任何信息, 这是否合适? 那么请问作者设置成人被试的原因是什么?

**回应:** 非常感谢您的问题! 本研究探讨幼儿对智能语音助手在不同领域的知识性信任发展情况, 以及智能语音助手的准确性对其是否有影响, 我们之所以设置成人被试, 是因为想将成人作为对照, 假设成人的发展为成熟的阶段, 那么幼儿的选择是否呈现出向成人靠近的趋势, 或者说幼儿的发展是否趋于成熟, 这是我们设置成人被试的初衷。根据您的建议, 我们已将成人被试的信息补充在摘要和实验设计中, 将成人被试的研究结果补充在讨论与结论中。

在摘要中改为“实验一, 90 名 4~6 岁幼儿和 30 名成人分别由不同类型信息源(智能语音助手 VS 人类)提供自然领域和社会领域的证言。结果显示, 幼儿与成人对智能语音助手的信任程度均较高, 成人对智能语音助手的信任程度高于幼儿。与社会领域的问题相比, 幼儿和成人更愿意向智能语音助手询问自然领域的问题。实验二以实验一为基础, 准确性条件进一步被操纵(智能语音助手准确, 人类不准确 VS 智能语音不准确, 人类准确)。在不同准确性条件下, 智能语音助手和人类分别向 90 名 3-5 岁幼儿和 30 名成人提供了关于自然与社会领域问题的证言。结果发现, 当智能语音助手不准确时, 幼儿和成人对语音助手的信任偏好均消失。”(p1)

在两个实验设计中分别改为“4(年龄: 4 岁/5 岁/6 岁/成人)”和“4(年龄: 3 岁/4 岁/5 岁/成人)”。(p8、p13)

在讨论中的补充有:

“同时, 我们可以看到 6 岁幼儿的决策正在向成人靠近, 趋于成熟。在操纵智能语音助手准确性的条件下, 实验二结果表明, 相较于 3 岁幼儿, 4 岁、5 岁幼儿和成人均更倾向于选择信任准确的智能语音助手, 在“小爱不准确, 圆圆准确”的条件下, 5 岁幼儿的决策已经与成人接近。”(p17)

“幼儿对智能语音助手的知识性信任存在领域特殊性。实验一结果表明, 相较于社会领域问题, 成人与幼儿在自然领域问题的初始询问中更倾向于选择智能语音助手, 5 岁、6 岁幼儿与成人在自然领域对智能语音助手的平均询问意愿、平均选择信任和平均最终判断均高于社会领域。……而在操纵智能语音助手和人类准确性的条件下, 成人与幼儿的选择均不存在领域特殊性, 在自然领域和社会领域中, 成人与幼儿更倾向于选择信任准确的信息提供者。这更加说明智能语音助手的准确性可能是幼儿对其知识性信任的先决条件, 准确性是幼儿更重视的好的信息提供者的特征。”(p18)

将结论改为:

“(1)幼儿对智能语音助手的知识性信任存在年龄差异, 年龄越大的幼儿越倾向于信任智

能语音助手。相对于幼儿，成人更愿意选择信任智能语音助手。

(2)准确性决定了幼儿和成人对智能语音助手的知识性信任，幼儿与成人仅对准确的智能语音助手存在知识性信任的偏好。

(3)幼儿与成人对智能语音助手的知识性信任存在领域特殊性，相较于社会领域问题，幼儿与成人更愿意向智能语音助手询问自然领域问题。”(p19)

再次感谢您的宝贵意见！

**意见 4:** 实验结果的呈现和分析应更为清晰和合理。

比如，在实验一的结果部分，作者做了三次重复测量的方差分析，第一次表明领域主效应和年龄主效应显著，第二和第三次都只显示年龄主效应显著。请问第二和第三次领域主效应如何？即便不显著，也是需要交代的，不应只字不提。同样地，实验二的实验设计提到有四个自变量，但是实验二多次重复测量方差分析结果的呈现，并未提及所有的自变量，这是否合适？而且，引言并未提及性别因素，实验设计中也并未将性别列为自变量，却在统计结果中突然偶尔地将性别作为自变量进行分析，这又是为何？

**回应:** 非常感谢您的建议！我们对统计部分做了修改，再利用 GLMM 模型进行的统计分析中，我们将涉及的自变量的统计结果全部报告，在实验一中包括年龄、领域的主效应和两者的交互效应，在实验二中包括准确性条件、年龄、领域的主效应和两者的两两交互效应，我们在您的意见“2.两个实验数据统计的方式令人疑惑”(修改意见回复 p10)中已详细列出我们修改过的统计结果，为免重复阐述，请您参阅，希望能更详细、全面地回答您的问题。

在此还需特别说明的是，在我们曾尝试将准确性条件、年龄、领域的三重交互纳入模型，但发现模型拟合度降低且部分模型无法拟合，因此并未将其纳入最终模型。此外，正如我们在您对实验二提出的疑惑 4 的回答中详细说明的，性别并不是我们的自变量，将其纳入随机变量也并未改善模型的拟合度，因此我们将性别作为无关变量，并未将其纳入统计分析中。

又如，“以上结果说明，不同领域问题影响幼儿初始信任智能语音助手的偏好，相较于社会领域问题，幼儿在自然领域问题的初始询问中更倾向选择智能语音助手。同时，不同年龄的幼儿对智能语音助手的初始询问偏好也不同，5 岁、6 岁幼儿更愿意选择智能语音助手作为信息提供者。而相较于幼儿，成人在初始询问中更愿意选择智能语音助手。”整段只提“幼儿”在不同领域中询问意愿的情况，不提成人在这两个领域的具体情况，那请问成人呢？就图 3 来看，成人在这两个领域对智能语音助手询问意愿的比例，都显著高于随机水平，并且在自然领域中的询问意愿程度似乎更高？既然成人也是被试，图 3 也呈现了成果的数据，在分析的时候为何有时候分析成人，有时候略过成人，是否有什么特别的考虑吗？

**回应:** 非常感谢您的宝贵建议！我们之所以设置成人被试，是因为想将成人作为对照，假设成人的发展为成熟的阶段，那么幼儿的选择是否呈现出向成人靠近的趋势，或者说幼儿的发展是否趋于成熟，这是我们设置成人被试的初衷。根据您的建议，我们在修改中将成人结果全部纳入分析对比。在修改稿中，我们在“2.6 结果与讨论”“3.6 结果与讨论”“4.讨论”中均补充成人的结果，详见 p9、p14，在对您的意见“3.研究被试群体和论文内容之间不匹配”的回答(修改意见回复 p16)中，我们也将我们的修改详细地列出，请您参阅。例如您在此处提到的“2.6.1 幼儿对信息提供者证言的询问意愿”中的结论部分，我们将其修改为“以上结果说明，成人与幼儿对智能语音助手的初始信任偏好受到不同问题领域的影响，相较于社会领域问题，成人与幼儿在自然领域问题的初始询问中均更倾向选择智能语音助手。同时，相较于幼儿，成人在初始询问中更愿意选择智能语音助手，年龄越大的幼儿越倾向于选择询问智能语音助手。”(p9)

此外，在实验一和实验二的结果部分，在没有文字引述的情况下，都直接先呈现了数据图。建议用文字引出数据图表，即先呈现文字，再呈现数字。

回应：非常感谢您的建议！根据您的建议，我们对实验一和实验二的结果部分已进行相应修改，都用文字引出数据图表。(p9、p14)

“对幼儿及成人在自然领域和社会领域对智能语音助手的平均询问意愿进行分析，结果如图 4 所示。”(p9)

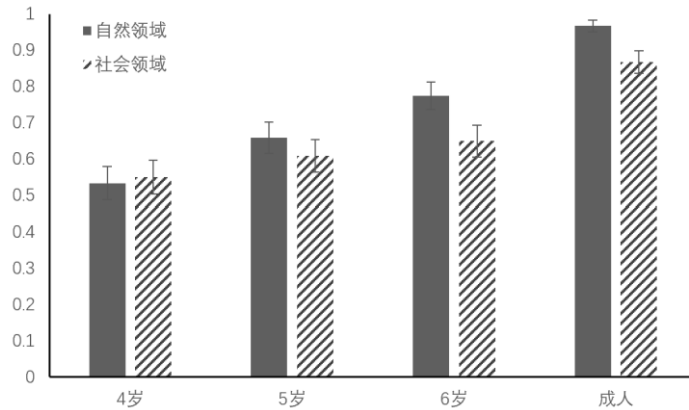


图 4 幼儿与成人在不同领域对智能语音助手的平均询问意愿

“对幼儿及成人在自然领域和社会领域对智能语音助手的平均外显信任进行分析，结果如图 5 所示。”(p10)

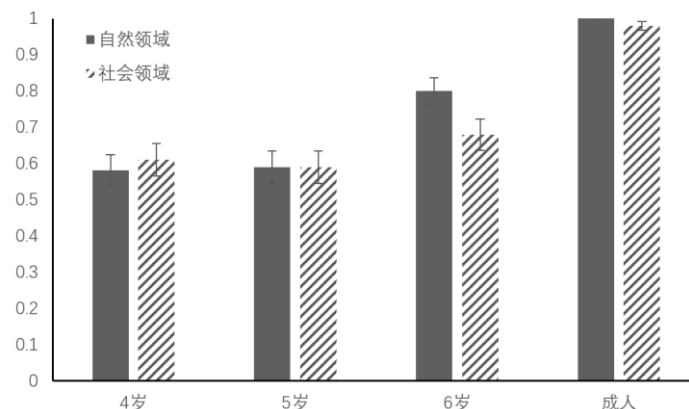


图 5 幼儿与成人在不同领域对智能语音助手的平均外显信任

“对幼儿及成人在自然领域和社会领域对智能语音助手的平均最终判断进行分析，结果如图 6 所示。”(p11)

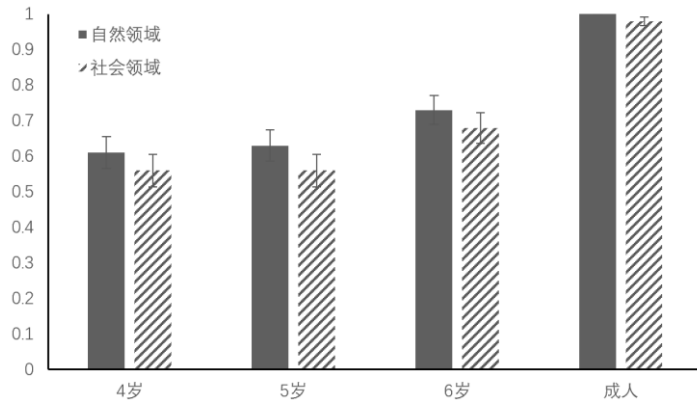


图6 幼儿与成人在不同领域对智能语音助手的平均最终判断

“在不同准确性条件下，对幼儿及成人在自然领域和社会领域对智能语音助手的平均询问意愿进行分析，结果如图8所示。当智能语音助手失去其准确性时，三个年龄段的幼儿及成人对智能语音助手在自然领域和社会领域的平均询问意愿均下降。”(p14)

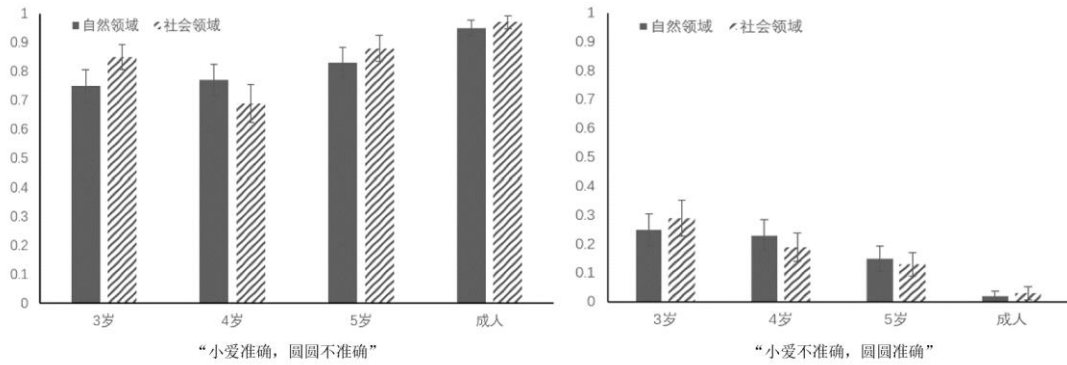


图8 不同准确性条件下，幼儿与成人在不同领域对智能语音助手的平均询问意愿

“.....准确性与年龄的交互作用如图9所示。.....”(p14)

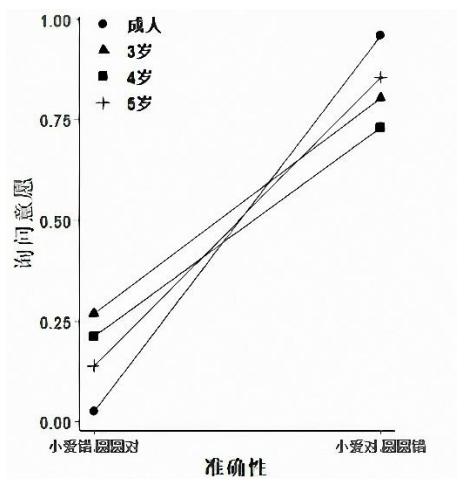


图9 准确性与年龄的交互作用

“在不同准确性条件下，对幼儿及成人在自然领域和社会领域对智能语音助手的平均外显信任进行分析，结果如图 10 所示。当智能语音助手失去其准确性时，三个年龄段的幼儿及成人对智能语音助手在自然领域和社会领域的平均外显信任均下降。”(p15)

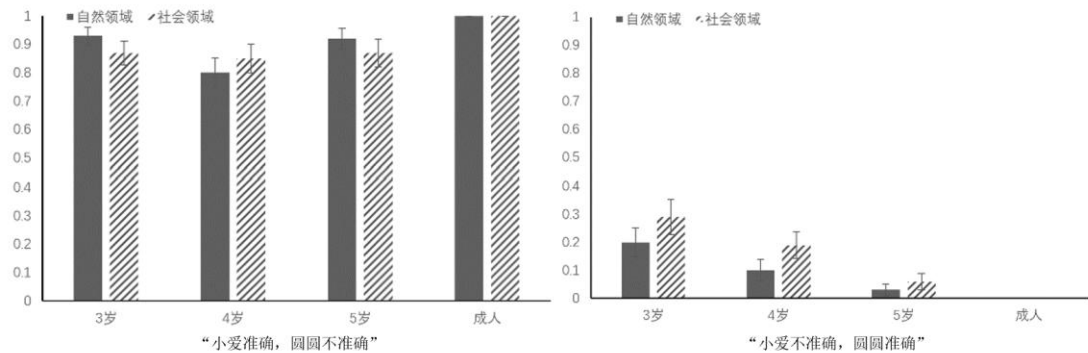


图 10 不同准确性条件下，幼儿与成人在不同领域对智能语音助手的平均外显信任

“在不同准确性条件下，对幼儿及成人在自然领域和社会领域对智能语音助手的平均最终判断进行分析，结果如图 11 所示。当智能语音助手失去其准确性时，三个年龄段的幼儿及成人对智能语音助手在自然领域和社会领域的平均最终判断均下降。”(p16)

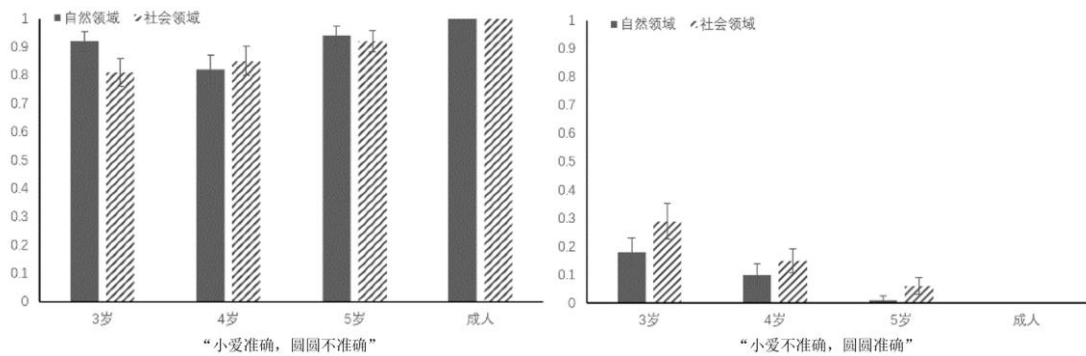


图 11 不同准确性条件下，幼儿与成人在不同领域对智能语音助手的平均最终判断

**意见 5:** 讨论和结论的准确性需要进一步斟酌。

比如，论文在讨论部分提及“幼儿对智能语音助手的知识性信任存在领域特殊性。”但是实际上，实验二似乎并没有延续这个特点。甚至，作者在将“问题的领域：自然 VS 社会”作为实验二自变量的情况下，在整个实验二结果的呈现和分析中，都没有提及“问题的领域”这一自变量对因变量的影响。在多个实验都涉及某个自变量的情况下，对该自变量如何影响因变量的解读，理应综合多个实验的结果，而不是选择性地采用某一实验的结果、忽视其他实验与此有关的结果。“幼儿对智能语音助手的知识性信任存在领域特殊性”这个论点之下，作者只用实验一的结果作为依据进行解读，只字不提“问题的领域”在实验二中的情况。如此，这个论点和后文的类似结论，是否就忽略了实验二的发现？那这个论点和后文类似结论成立的条件是否也同样被忽视了？那么，这个论点和后文类似结论是否可信？

**回应:** 非常感谢您的宝贵意见！诚如您所说，我们应综合多个实验的结果，而不应忽视实验二中领域相关的结果，我们在“幼儿对智能语音助手的知识性信任存在领域特殊性。”这部分的讨论中补充了实验二的相关内容：“而在操纵智能语音助手和人类准确性的条件下，成人与幼儿的选择均不存在领域特殊性，在自然领域和社会领域中，成人与幼儿更倾向于选择信任准确的信息提供者。这更加说明智能语音助手的准确性可能是幼儿对其知识性信任的先决



条件，准确性是幼儿更重视的好的信息提供者的特征。”(p18)

在“儿童对智能语音助手的知识性信任存在年龄差异。”的讨论中也增加了实验二的相关内容：“在操纵智能语音助手准确性的条件下，实验二结果表明，相较于3岁幼儿，4岁、5岁幼儿和成人均更倾向于选择信任准确的智能语音助手，在‘小爱不准确，圆圆准确’的条件下，5岁幼儿的决策已经与成人接近。”(p17)

**意见 6：摘要需要优化。**

建议参考《心理学报》投稿指南“中文摘要 300 字左右”和“尽可能不用‘本研究’、‘我们’等第一人称”“所有数字和英文字体都用 Times New Roman”等要求，适当精简和优化中文摘要。

**回应：**非常感谢您的建议！我们已按照要求修改摘要为：

“随着智能语音技术的不断成熟，智能语音助手已经成为数字时代幼儿的重要信息提供者。本研究采用双信息源范式，考察个体对智能语音助手知识性信任的发展，并验证准确性对幼儿评估智能语音助手可靠性的作用。实验一，90名4~6岁幼儿和30名成人分别由不同类型信息源(智能语音助手 VS 人类)提供自然领域和社会领域的证言。结果显示，幼儿与成人对智能语音助手的信任程度均较高，成人对智能语音助手的信任程度高于幼儿。与社会领域的问题相比，幼儿和成人更愿意向智能语音助手询问自然领域的问题。实验二以实验一为基础，准确性条件进一步被操纵(智能语音助手准确，人类不准确 VS 智能语音不准确，人类准确)。在不同准确性条件下，智能语音助手和人类分别向90名3-5岁幼儿和30名成人提供了关于自然与社会领域问题的证言。结果发现，当智能语音助手不准确时，幼儿和成人对语音助手的信任偏好均消失。综上所述，智能语音助手是3~6岁幼儿获取信息的主要来源，信息提供者的类型、信息提供者的准确性以及知识领域都会影响幼儿的知识性信任。幼儿对智能语音助手知识性信任存在领域特殊性，准确性是衡量智能语音助手可靠性的重要标准。”(p1)

**意见 7：逻辑需要进一步斟酌，论文在没有转折含义之处用了转折关联词。**

比如，“而在社会领域中，成人组选择询问智能语音助手的概率与随机概率相比差异显著( $\chi^2 = 64.53, df = 1, p < 0.001$ )；而幼儿组中5岁、6岁幼儿选择询问智能语音助手的概率与随机概率的差异显著(5岁组 $\chi^2 = 5.63, df = 1, p = .018$ ;6岁组 $\chi^2 = 10.80, df = 1, p = .001$ )”。就数据和前后文来看，成人、5岁和6岁幼儿在社会领域中询问智能语音助手的情况，跟他们在自然领域中的情况，具有相似性，都显著高于随机水平，似乎并没有出现前后转折。后文还有类似的情况，如“而在社会领域中，成人组全部选择信任智能语音助手(100%)”。

**回应：**非常感谢您的宝贵建议！由于我们更改了统计方式，此部分文字已被删除。根据您的建议，我们再次检查全文段落、语句中的逻辑关系，对使用不当的关联词进行修改。

**意见 8：建议进一步规范论文的格式。**

**问题 1：**文中参考文献的格式不够规范，比如，请核对文中括号里年份跟逗号之间的间隔是否符合要求。

**问题 2：**文中括号里多个参考文献中的分号，有时跟前面的内容之间有空格，有时没有空格，如(Corriveau et al.,2014; Einav et al.,2013; Clément et al.,2004;张璟等,2020)。

**问题 3：**文中一些括号是宋体格式，一些是 Times New Roman 格式。

**问题 4：**一些标题的序号和文字之间有空格，一些没有。

**问题 5：**文中本不该用半格的标点符号，用了半格。如“，成人更愿意选择询问智能语音助手”。

问题 6: 不该少半边括号的情况, 少了半边括号。如“2 问题的领域: 自然 VS 社会)”。

问题 7: 各图表序号和文字之间的间隔是否符合要求?

问题 8: 正文中个别句子作为首行, 本该缩进两个汉字字符的宽度, 但是没有缩进。

问题 9: 正文中有时是“ $p<.001$ ”, 有时是“ $p<0.001$ ”。

问题 10: “(成人-4 岁:  $p<0.001$ ; 成人-5 岁:  $p<0.001$ ; 成人-6 岁:  $p<0.001$ ); 幼儿组中 5 岁、6 岁幼儿选择智能语音助手的询问意愿显著高于 4 岁幼儿(6 岁-4 岁:  $p<0.001$ , 5 岁-4 岁:  $p=.043$ )”。冒号后面紧跟逗号是啥意思?

问题 11: 前后表述不一致, 如前两次为“Siri”, 第三次为“siri”。不仅首字母不同, 字体格式也不同了。

总之, 请注意逐一检查全文的格式, 提高规范性。

回应: 非常感谢您的建议! 十分惭愧, 我们已经仔细检查全文的格式问题, 并逐一修改, 在此对您列举的问题做出回复:

问题 1、2: 根据您的建议, 我们再次仔细阅读《心理学报》的“参考文献著录要求”, 规范正文中的文献引用格式, 逗号与年份之间空一格, 在引用多个文献的情况下, 分号与作者之间空一格。同时规范文后文献列表中的格式, 英文逗号、点号、冒号后均留一空格。

问题 3: 根据您的建议, 我们已将文中所有括号格式改为 Times New Roman 格式。

问题 4: 根据您的建议, 我们统一全文标题格式, 在标题的序号和文字之间空一格。

问题 5、10: 根据您的建议, 我们仔细检查全文的标点符号, 已修改文章中使用不当的标点符号。

问题 6: 根据您的建议, 我们仔细检查全文括号的使用, 已逐一修改。

问题 7: 根据您的建议, 我们按照在图表序号与文字之间空一格。

问题 8: 根据您的建议, 我们仔细检查全文各段落的首行缩进, 已逐一修改。

问题 9: 根据您的建议, 我们将所有统计表达统一为 0.001。

问题 11: 根据您的建议, 我们已将其统一为“Siri”。

再次感谢您指出的不妥之处!

意见 9: 建议进一步优化句子的表述, 提高论文的可读性。

文中有问题的句子比较多, 全文读起来晦涩、不痛快, 论文可读性亟待提升。建议: 加强挑剔性阅读的工作, 厘清每个句子各部分的逻辑; 适当使用标点符号; 消除语法问题; 更正打印错误; 准确地使用助词“的、地、得”。

回应: 非常感谢您的宝贵建议! 深感惭愧, 根据您的建议, 我们仔细阅读全文, 逐一修改存在问题的地方, 理顺每个句子各部分的逻辑, 已修改标点符号的误用、语法问题、打印错误、助词的误用等错误, 以加强文章的可读性。

例 1, “以往的研究主要集中在幼儿对以人为信息源的认识性与社会性因素差异为主的知识性信任发展上, 对于考量幼儿对智能语音系统知识性信任发展的研究非常匮乏, 并且就研究者所知, 目前仅有一项研究直接对比了幼儿对智能语音系统和人类的知识性信任差异(Girouard-Hallam & Danovitch,2022)。”

回应: 我们已修改为: “以往研究多侧重于探究幼儿对人类信息源的知识性信任发展情况及影响因素, 考察幼儿对智能语音系统的知识性信任发展的研究非常匮乏, 就研究者所知, 目前仅有一项研究直接对比了幼儿对智能语音系统和人类的知识性信任差异(Girouard-Hallam & Danovitch, 2022)。”(p2)

例 2, “同时, 研究者对依附于 AI 智能系统的机器人作为信息提供者的幼儿知识性信任

展开了较多的研究。”这样的句子，重点不突出，读起来也比较拗口。或许可以考虑改为“同时，幼儿对 AI 智能系统机器人的知识性信任，颇受研究者关注。”

回应：非常感谢您的建议！为使文章逻辑更加通顺，此部分文字已被删除。

例 3，比如，“近年来随着智能语音技术的不断进步和产品的普及，儿童接触各类智能语音助手的机会和时间大大增加，关于儿童如何对智能语音助手提供的信息进行选择性判断和采信的研究不但有助于我们对儿童基于智能语音助手学习的机制的理解，也是当前社会的广泛需求。”

回应：为使文章逻辑更加通顺，此部分文字已被删除。

例 4，“结果表明即使是 3 岁的幼儿也已经能够根据信息提供者所提供线索的准确性作为依据进行知识性信任的判断，”

回应：我们已修改为：“结果表明，即使是 3 岁的幼儿，也能根据信息提供者提供的信息的准确性进行知识性信任的判断，”(p18)

例 5，“未来研究可收集幼儿是否认为智能语音助手具有教学能力等情感态度以进一步探究其对智能语音助手知识性信任发展。”

回应：由于我们修改了讨论部分的有关内容，为保证文章的逻辑性，此句已被删除。

例 6，“三个年龄组的幼儿以及成人在自然与社会领域下选择信任智能语音助手或人类的概率与随机概率（50%）的差异情况如图 3 所示。”

回应：由于我们修改了统计方式，此句已被删除。

例 7，“即更容易不加选择的相信他人他们被告知的内容”。

回应：我们已将其修改为：“即更容易不加选择地相信他人提供的信息。”(p2)

例 8，“其结果将有利的证明儿童对智能语音助手的知识性信任是否仅基于其准确性。”

回应：我们已将其修改为：“研究结果将验证儿童对智能语音助手的知识性信任是否取决于其准确性。”(p3)

例 9，“然而这项研究的被试年龄跨度较大，且不是连续的，对幼儿对智能语音助手的知识性信任的发展情况的考察存在一定的不完整性，特别考虑到 4 岁以下幼儿可能存在的轻信倾向，有必要通过从更年幼儿童开始连续考察幼儿期儿童对智能语音助手的知识性信任的发展情况。”

回应：我们已将其修改为：“然而这项研究的被试年龄跨度较大，存在一定程度的非连续性和不完整性，特别是考虑到 4 岁以下幼儿可能存在的轻信倾向，因此有必要从更年幼的儿童开始，连续考察学前期儿童对智能语音助手的知识性信任的发展情况。”(p2)

例 10，“考虑信息源的知识性状态专业性和准确性”“结果发现随年龄增长儿童越来越倾向于从语音助手寻求事实信息。”

回应：我们已将其修改为：“考虑信息源的专业性与准确性”(p3) “结果发现，随着年龄的增长，儿童更倾向于向智能语音助手寻求事实信息。”(p2)

例 11，“因为智能语音助手并不具备的社会联系、道德和动机等，在原则上，可能仅有

一个标准可以同样适用于智能语音助手，即准确性。”

回应：我们已将其修改为：“因为智能语音助手不具有道德、动机等属性，原则上可能只有一种标准能够适用于智能语音助手，即准确性。”(p3)

例 12，“倾向于从信任之前准确的电脑而非不准确的电脑呈现的信息”

回应：我们已将其修改为：“倾向于信任以前准确的电脑提供的信息，而不相信错误的电脑。”(p3)

例 13，“张璟等(2020)发现在线索冲突情境下幼儿对自然与社会领域知识的选择性信任具有相同的年龄发展趋势，”

回应：我们已将其修改为：“张璟等(2020)发现，在线索冲突情境下，幼儿对自然和社会领域知识的选择性信任具有不同的年龄发展特点。”(p3)

例 14，“此次实验没有在前测或是后测幼儿家中拥有智能语音助手的情况，”

回应：我们已将其修改为：“本研究没有对幼儿家中是否拥有智能语音助手进行前测或后测，”(p18)

例 15，甚至在论文的结论部分，也存在类似的情况：

“（1）幼儿对智能语音助手的知识性信任存在年龄差异，随年龄增长，表现出与人类相比对智能语音助手的知识性信任偏好。。

（2）智能语音助手的准确性决定了幼儿对其知识性信任，幼儿仅对准确的智能语音助手存在知识性信任的偏好。。”

回应：我们已将其修改为：

“(1)幼儿对智能语音助手的知识性信任存在年龄差异，年龄越大的幼儿越倾向于信任智能语音助手。相对于幼儿，成人更愿意选择信任智能语音助手。

(2)准确性决定了幼儿和成人对智能语音助手的知识性信任，幼儿与成人仅对准确的智能语音助手存在知识性信任的偏好。

(3)幼儿与成人对智能语音助手的知识性信任存在领域特殊性，相较于社会领域问题，幼儿与成人更愿意向智能语音助手询问自然领域问题。”(p19)

类似情况实在是多，就不逐一列举了，请注意核查。

回应：再次感谢您指出的问题！很抱歉我们在之前的文章中出现格式不规范的地方，我们已逐一检查、修改，非常感谢！

意见 10：个别概念的使用需要再斟酌，个别字母缩写指代的含义需补充说明。

例 1，“幼儿”在中文语境中，通常是指 3-6 岁的孩子，因此“7~8 岁幼儿”的表述，改为“7~8 岁儿童”是否更合适？

例 2，在同一段落中（如“热身阶段”），交流的环节用“儿童”指代被试，叙述指导语的时候用“幼儿”指代被试。建议尽量保持一致。

例 3，“有关智能语音助手作为信息源的研究中发现准确性对幼儿对智能语音助手的知识性信任可能有较大影响，如有研究表明幼儿会要求 DVAs 在回答问题前证明其准确性 (Yarosh et al.,2018)。”建议用括号简要标注说明何为“DVAs”。

回应：非常感谢您的建议！我们已逐一进行修改：

例 1，根据您的建议，已改为“7~8 岁儿童”。(p2、p3)

例 2, 根据您的建议, 我们已统一用“幼儿”指代被试。(p7)

例 3, 根据您的建议, 同时为保证上下文一致, 我们已将“DVAs”改为“智能语音助手”。  
(p3)

意见 11: 建议进一步优化文中所有图的清晰度。

回应: 非常感谢您的建议, 我们已优化文中所有图的清晰度。

综合考虑上述所有问题, 建议修改后再审。

回应: 非常感谢您的上述宝贵建议! 您的意见和建议给我们很大的启发, 我们已根据您的和另一位审稿专家的建议逐一修改全文, 再次感谢您的宝贵建议!

---

## 第二轮

审稿人 1 意见:

意见 1: 我的问题还是关于研究意义, 虽然作者陈述, “有研究表明, 幼儿能够基于准确性对电子媒体信息源(如电脑等)进行知识性信任, 但就我们所知, 目前没有研究直接探讨准确性是否影响幼儿对智能语音助手的知识性信任”。电子媒体信息源(如电脑等)与智能语音助手在本质上有什么区别?

回应: 非常感谢您的宝贵意见! 我们之前的表述可能不够清楚, 根据您的意见, 我们已在正文中进一步说明了电子媒体信息源(电脑等)与智能语音助手在本质上的区别(P3), 在此对您的问题进行了回答。

电子媒体信息源(电脑等)与智能语音助手最核心的区别在于, 人类与智能语音助手的互动方式为半社交型, 而人类与传统电子媒体(电脑等)的互动方式大多为非社交型。

首先, 传统电子媒体(电脑等)针对人类的特定查询(如通过键盘打字输入), 从庞大的知识库(如因特网)中选择信息并将其以链接的形式呈现给用户(Noles et al., 2015)。因此, 传统电子媒体(电脑等)并不具备社会属性。然而, 智能语音助手作为基于互联网的设备, 通过聆听唤醒词(如“Hi Siri”、“Hey Google”)、记录语音输入并发送信息来回应人类的语音(Murray, 2021), 具有人类的特殊品质——语音。智能语音助手的输出是根据幼儿“询问”而提供的, 幼儿可以通过调整“询问”内容与智能语音助手进行语音交互。

其次, 智能语音助手可以为幼儿提供互动式学习问答与情感交流, 而传统电子媒体不具有这样的功能。随着人工智能技术的进步, 智能语音助手具备自然语言处理等理解能力, 能够通过自然聊天的方式识别幼儿的意图并给出适当的答复与建议。同时, 随着智能语音助手开放域情感性的增强(例如基于人类反馈的强化学习技术 (Reinforcement Learning with Human Feedback) 的 DVA(如 Chat Generative Pre-trained Transformer, ChatGPT)), 智能语音助手能够根据聊天环境学习和理解人类语言, 增强与幼儿的情感互动。

最后, 幼儿对两者的归因(attribution)也存在明显区别。有研究证明, 超多半数的 3~10 岁儿童认为智能语音助手具有部分心理属性(聪明的等)(92.80%)、道德属性(能和“我”成为朋友等)(65.00%)和社会属性(能教“我”知识等)(89%) (Xu & Warschauer, 2020; Girouard-Hallam et al., 2021; Girouard-Hallam & Danovitch, 2022a)(p3)。与之相反, 有研究表明, 5 岁幼儿认为电脑并没有生命, 也不具备情感(Scaife & Van Duuren, 1995)。

因此, 智能语音助手是同时具有社会属性(例如, 能用语言与人类交流)和非社会属性(例如, 缺乏非语言线索: 点头、面部表情等)的半社会型媒体(Girouard-Hallam et al., 2021), 不

属于现有的任何本体论类别(Kahn et al., 2013), 也与传统电子媒体也有着本质上的区别(p3)。考虑到以往研究多探究不具备社会属性的电脑等传统电子媒体的准确性对幼儿知识性信任的影响, 并没有研究直接探讨具有半社会属性的智能语音助手的准确性是否影响幼儿对其的知识性信任。因此, 智能语音助手独特的半社会属性决定了探究幼儿对智能语音助手的知识性信任具有独特的理论价值与现实价值。

**意见 2:** 作者指出 Girouard-Hallam 和 Danovitch(2022)的研究范式的局限性, 在我看来 Girouard-Hallam 实验二已经消除了 yes bias 的问题, 本研究的操作跟其类似, 如何理解作者说的“但这种修改并未从根本上弥补该研究方法的缺陷”。

**回应:** 非常感谢您的宝贵意见! 我们之前的表述可能不够清楚, 根据您的建议, 我们已在正文中进行相应的修改和补充(p2), 在此先对您的问题回答。

Girouard-Hallam 和 Danovitch(2022)在实验一中采用单一信息源范式, 主试提出问题, 由智能语音助手或人类回答问题, 然后让幼儿判断智能语音助手或人类的回答是否正确, 该范式可能存在轻信偏见的问题。在其研究的实验二中, Girouard-Hallam 和 Danovitch(2022)将实验范式修改为双信息源范式, 在屏幕中央显示问题图片, 左右两边呈现智能语音助手和人类的照片, 主试提出问题, 并让幼儿回答“谁能给我们正确的答案?”, 该范式虽然弥补了轻信偏见的缺陷, 但仍存在的以下两个主要的问题:

第一, Girouard-Hallam 和 Danovitch(2022)的实验二的信息提供者并未提供证言信息。知识性信任(epistemic trust)是指儿童对信息源提供信息(即证言, testimony)的信任过程(张耀华, 朱莉琪, 2014)。儿童的学习过程被概念化为对信息源提供信息的判断、选择和采信过程(张耀华 & 朱莉琪, 2014)。然而, 在 Girouard-Hallam 和 Danovitch(2022)的实验二中, 作为信息提供者的智能语音助手和人类均没有提供证言, 在主试提出问题后, 两名信息提供者并未回答问题, 主试直接向幼儿询问“谁能给我们正确的答案?”, 并且在整个实验过程中, 信息提供者均未提供证言。也就是说, 幼儿在实验过程中并没有和智能语音助手及人类进行有效的选择性社会学习互动, 幼儿没有经历判断、选择和采信证言的过程, 严格意义上来说并没有进行完整的知识性信任。

第二, Girouard-Hallam 和 Danovitch(2022)的实验二中, 仅设置了有关知识性信任的询问问题:“谁能给我们正确的答案?”以往大量研究发现, 询问问题并不是知识性信任的可靠指标, 也不能体现幼儿的知识性信任过程。幼儿此时的回答可能仅仅是表达选择询问谁的意愿, 并不代表对信息提供者的知识性信任因此在本研究中增加了外显信任和最终判断问题, 直接考察幼儿对双方证言的评估, 及对信息提供者可信性的推断(李婷玉 等, 2018), 完整考察幼儿是否真正采信了智能语音助手提供的证言。

由此, 本研究认为 Girouard-Hallam 和 Danovitch(2022)的实验二虽然弥补了轻信偏见的缺陷, 但仍存在信息提供者并未提供信息、知识性信任实验任务不完整等方面的不足。因此, 本研究将采用知识性信任研究的双信息源范式, 并设置询问意愿、外显信任和最终判断的任务, 完整考察 3~6 岁幼儿对智能语音助手的知识性信任发展。

**意见 3:** 小修改, P3“在与时间有关的话题中, 6 岁幼儿认识到触摸屏对最新信息的效用, 他们倾向于向触摸屏询问今天的天气, 而选择在书本中了解昨天的篮球比赛。”书本中了解昨天的球赛, 这个例子不恰当, 这些信息通过电子媒体也能了解。

**回应:** 非常感谢您的宝贵意见! 我们已删除与这个例子相关的内容。

.....

**审稿人 2 意见:** 作者对论文中此前存在的问题进行了认真修改, 论文质量得以实质提升。目

前，还有两个地方需要注意。

**意见 1:** 一是中文摘要适当再精简，以进一步契合《心理学报》的要求。

**回应:** 非常感谢您的建议！根据您的建议，我们对中文摘要进一步修改，已经符合学报的中文摘要字数要求：

“基于双信息源实验范式，考察 3~6 岁幼儿和成人对智能语音助手(Digital Voice Assistants, DVAs)的知识性信任，并验证准确性在其信任中的作用。实验一，操纵了信息源类型(DVAs VS 人类)和知识领域(自然 VS 社会)。结果显示，与人类相比，幼儿和成人对 DVAs 的知识性信任更强。对 DVAs 的信任程度，6 岁幼儿高于 4、5 岁幼儿，成人高于幼儿。与社会领域相比，幼儿和成人均更倾向于询问 DVAs 自然领域的问题。基于实验一，实验二操纵了两种信息源提供证言的准确性条件，结果发现，当 DVAs 不准确时，幼儿和成人均不再信任其证言，且 3 岁幼儿对准确性的敏感性较低。研究结果表明，信息源类型、知识领域和证言准确性都会影响幼儿的知识性信任，准确性是评估 DVAs 可靠性的重要标准。”(p1)

**意见 2:** 二是文后参考文献的格式，仍然需要对照《心理学报》的格式和要求，逐一检查、规范。

**回应:** 非常感谢您的建议！根据您的建议，我们已经按照对《心理学报》的格式和要求对参考文献格式逐一检查、修改，例如，英文点号和逗号后，多留一空格；文章题目和书名，只有第一个单词的首字母需大写，其余为小写，特殊要求大写的单词除外等。

---

### 第三轮

**编委意见:**

据两位审稿人意见，作者做了认真修改，但是目前仍然有一些语言、格式、特别是英文摘要等问题，希望认真修改！

**回应:** 感谢编辑委员专家提出的宝贵意见。根据编委专家的意见，我们对这篇文章的语言和格式作了全面的检查和修改。

**语言问题:** 我们修改了文章中仍然存在的语法和句子表达不清等问题，例如，将“考虑到 DVAs 与传统电子媒体(电脑和网页)相比，可与幼儿进行语音交互，为幼儿提供互动式学习问答与情感交流，对幼儿更具吸引力。”修改为“考虑到与传统电子媒体(电脑和网页)相比，DVAs 可与幼儿进行语音交互，能够为幼儿提供互动式的学习问答与情感交流，对幼儿具有更强吸引力。”(p3)

**格式问题:** 我们再次检查了全文的格式，并根据心理学报投稿指南的要求进行了逐一修改。例如，根据“希腊字母( $\alpha$ 、 $\beta$ 、 $\chi^2$ 、 $\eta^2$ )、上标和下标的数字和文字，不用斜体。”的要求，将文中的“ $\chi^2$ ”全部修改为“ $\chi^2$ ”；根据“英文题目——四号”的要求，将英文题目修改为四号 Times New Roman 格式。

**英文摘要问题:** 我们已经请英语专业人士（英语翻译专业的副教授）对本文的英文摘要进行了仔细的修改和润色。

---

### 第四轮

**主编意见:** 文章有创新性贡献，有几个小地方请补充或澄清：

回应：非常感谢主编的宝贵意见，我们已按照您的意见在文中进行了相应补充与说明。

意见 1：完善两个实验的伦理批件说明和幼儿及监护人知情同意。

回应：非常感谢主编的宝贵意见，我们已在正文中进行相应补充：“本实验已通过所在单位的伦理审核，并经幼儿所在幼儿园同意，由园方与幼儿的监护人进行沟通，已获得其监护人的知情同意。”(p4、12)

本研究的伦理审核已发送至《心理学报》编辑部邮箱。需要特别说明的是，我们招募被试的幼儿园为本单位附属幼儿园，所有被试幼儿的监护人在入学时与幼儿园统一签署了来自本单位科研项目的知情协议，本研究已征得园方的同意，已由园方与幼儿监护人进行了沟通，已获得监护人的知情同意。

意见 2：作 t 检验的地方，建议补充 Cohen'd 值。

回应：非常感谢主编的宝贵意见，我们在对数据进行事后多重比较的过程中使用了调整后的 t 检验，根据您的建议，我们已经在文章中所有作 t 检验的部分均补充了 Cohen'd 值。

实验一中，在“2.6.1 幼儿对信息提供者证言的询问意愿”中“对年龄进一步进行事后对比发现，相较于幼儿，成人更愿意选择询问 DVAs(成人-4 岁:  $t = 9.39, p < 0.001, d = 0.93$ ; 成人-5 岁:  $t = 7.10, p < 0.001, d = 0.72$ ; 成人-6 岁:  $t = 5.11, p < 0.001, d = 0.54$ )，年龄越大的幼儿越倾向于选择询问 DVAs(5 岁-4 岁:  $t = 2.30, p = 0.022 < 0.05, d = 0.19$ ; 6 岁-4 岁:  $t = 4.28, p < 0.001, d = 0.36$ ; 6 岁-5 岁:  $t = 1.98, p = 0.048 < 0.05, d = 0.17$ )。”(p9); 在“2.6.2 幼儿对信息提供者的外显信任”中“对年龄进一步进行事后对比发现，相较于幼儿，成人更愿意选择信任 DVAs(成人-4 岁:  $t = 10.48, p < 0.001, d = 1.12$ ; 成人-5 岁:  $t = 10.59, p < 0.001, d = 1.13$ ; 成人-6 岁:  $t = 6.62, p < 0.001, d = 0.79$ ); 相较于 4、5 岁幼儿，6 岁幼儿更愿意选择信任 DVAs(6 岁-4 岁:  $t = 3.86, p < 0.001, d = 0.31$ ; 6 岁-5 岁:  $t = 3.97, p < 0.001, d = 0.32$ )。”(p10); 在“2.6.3 幼儿对信息提供者的最终判断”中“对年龄进一步进行事后对比发现，相较于幼儿，成人更愿意最终信任 DVAs(成人-4 岁:  $t = 10.67, p < 0.001, d = 1.15$ ; 成人-5 岁:  $t = 10.46, p < 0.001, d = 1.13$ ; 成人-6 岁:  $t = 7.41, p < 0.001, d = 0.86$ ); 相较于 4、5 岁幼儿，6 岁幼儿更愿意最终信任 DVAs(6 岁-4 岁:  $t = 3.27, p = 0.001, d = 0.26$ ; 6 岁-5 岁:  $t = 3.05, p = 0.002 < 0.01, d = 0.25$ )。”(p11)

实验二中，在“3.6.1 幼儿对信息提供者证言的询问意愿”中“对准确性条件进行事后对比发现，幼儿更愿意选择询问准确的信息提供者( $t = 28.62, p < 0.001, d = 1.85$ )，不考虑是人还是 DVAs。”(p15); 在“3.6.2 幼儿对信息提供者的外显信任”中“对年龄进一步进行事后比较发现，相较于 3 岁幼儿，4 岁、5 岁幼儿和成人均更倾向于选择信任准确的信息提供者(4 岁-3 岁:  $t = 2.85, p = 0.004 < 0.01, d = 0.26$ ; 5 岁-3 岁:  $t = 3.95, p < 0.001, d = 0.36$ ; 成人-3 岁:  $t = 2.02, p = 0.043 < 0.05, d = 0.18$ )。对准确性条件进行事后比较发现，幼儿更倾向于选择信任准确的信息提供者( $t = 41.27, p < 0.001, d = 2.67$ )。”(p15); 在“3.6.3 幼儿对信息提供者的最终采信”中“对准确性条件进行事后比较发现，幼儿更倾向于最终信任准确的信息提供者( $t = 42.79, p < 0.001, d = 2.76$ )。”(p17)

意见 3：“年龄  $M = 262.22$ ”有误

回应：衷心感谢您提出的宝贵意见，我们为自己的疏忽深表歉意，已经修改为“年龄  $M = 21.85$ ”(p4)和“年龄  $M = 21.70$ ”(p12)。