寻求支援：儿童如何应对接连发生的机器人通信故障

Maria Teresa Parreira*
mb2554@cornell.edu 康奈尔大学 纽约，美国

Isabel Neto*
aineto@fc.ul.pt Lasige，里斯本大学科学学院，葡萄牙

Filipa Rocha Lasige，里斯本大学科学学院，ITI/LarSys高等技术学院，葡萄牙

Wendy Ju 康奈尔大学、康奈尔科技 纽约，美国

arXiv:2601.00754v1 [cs.RO] 2 Jan 2026

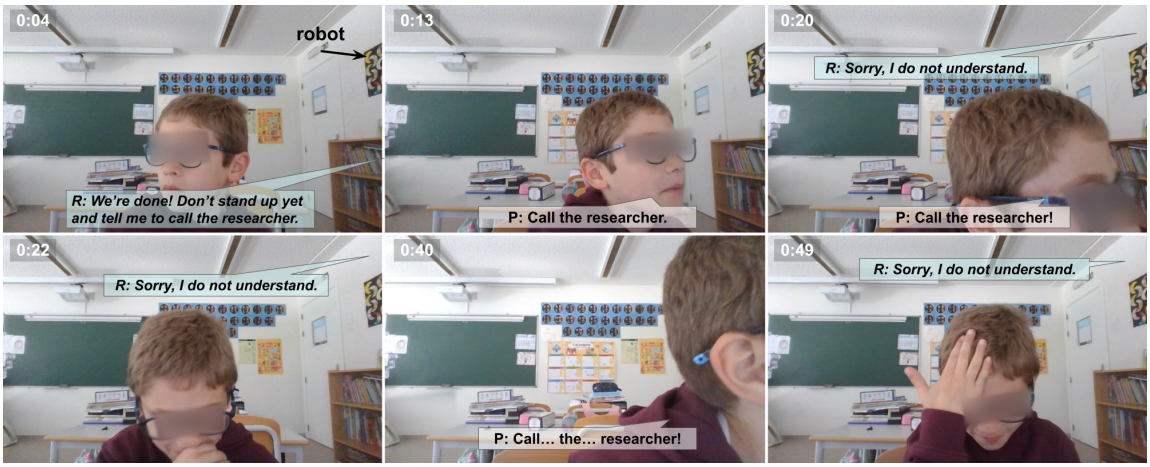

图 1：面对接连发生的机器人错误时的反应。给儿童下达指令后，机器人连续三次未能理解儿童的提示。在这个情境中，参与者通过靠近机器人并采用言语调整（语气强硬、放慢语速节奏）来应对故障。第三次出错后，儿童明显表现出沮丧和幽默，随后再次重复指令。之后，研究人员进入房间。补充数据可在研究资料库中找到。

摘要

儿童如何应对反复出现的机器人错误？虽然以往研究考察了成年人对连续机器人错误的反应，但儿童的反应仍鲜少受到关注。本研究探讨了儿童对机器人的社交错误和性能错误的反应。针对后者，本研究让儿童参与者（N=59，年龄为8-10 岁）参与 Liu 等人的连续机器人故障范式，以考察年轻用户如何应对机器人反复出现的对话错误。参与者与机器人互动时，机器人连续三次未能理解他们的提示语；研究人员对他们的行为反应进行了录像和分析。与原研究中的成人反应相比，我们发现了相似之处，也发现了差异。与成年人一样，儿童会调整提示语、改变言语语气，并在连续发生错误的过程中表现出愈发强烈的非语言情绪反应。然而，儿童表现出更多脱离行为，包括暂时不理会机器人或主动寻求成年人的帮助。错误并未影响参与者对机器人的看法，这表明儿童的对话预期更加灵活。这些发现为为年轻用户设计更有效且符合其发展阶段的机器人互动系统。

关键词

连续错误；机器人错误；可复现性；儿童—机器人互动；错误恢复；性能错误；社交错误

1 引言

机器人系统日益广泛地应用于各个领域，从教育陪伴机器人到医疗助理机器人，其用户群体已扩大到典型成年用户群体之外的人群 [5, 10, 16, 31]。其中包括儿童，与成年用户相比，他们可能具有不同的交互预期、沟通模式和错误容忍度 [13, 24, 32, 35]。

尽管机器人能力不断提升，但机器人仍不完美，在互动过程中难免会犯错。这些错误对用户的影响因用户和机器人的特征以及情境因素而异 [21, 22, 47]。尽管此前的人机交互（HRI）研究考察了成年人如何应对机器人故障 [7, 20, 27, 40, 47], ，但在理解不同用户群体，尤其是儿童，如何应对并从机器人错误中恢复方面，仍存在一个关键的

*作者对本研究的贡献相同。

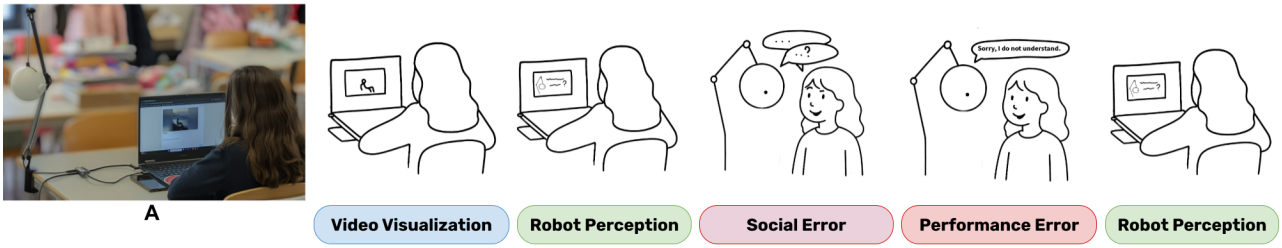

图2：实验设置（A）和研究流程。在简要介绍机器人 Simon 后，儿童会观看笔记本电脑上的一系列视频。随后，他们填写一份简短问卷，对机器人的不同感知维度进行评估。完成后，Simon 会向他们提问：根据实验条件，Simon 会打断他们的回答（社交错误），也可能不会。之后，Simon 无法理解参与者要求呼叫研究人员的请求（性能错误）。连续出现 3 次错误后，研究人员会回到房间。结束实验前，儿童会再次填写机器人感知问卷。

知识空白。连续机器人故障（即机器人接连多次发生故障）是一个尤具挑战性的情境：它在现实世界部署中十分常见，却鲜少得到研究。此外，对话中不恰当地打断他人等社交错误，会影响用户对机器人好感度的感知 [1]。了解儿童具体如何应对性能故障和社交违规，对于设计适合其年龄的错误检测和恢复策略至关重要。

Liu 等人 [27] 近期的研究探讨了成年人对机器人连续对话失败的反应，揭示了复杂的行为模式，包括沟通策略调整、情绪反应以及最终放弃互动。然而，儿童的认知、情绪和社会发展与成年人不同 [23, 32], ，可能因此表现出不同的反应模式，需要单独进行研究。

在本研究中，我们旨在探究“儿童如何感知并应对反复出现的机器人错误？”。为此，我们在儿童参与者中复现了连续机器人失败范式，并额外引入社交错误条件，以考察年幼用户对不恰当的机器人打断以及反复出现的机器人对话失败的反应。通过视频分析和行为编码，我们刻画了儿童的反应，并将其与成人反应进行比较，为开发更稳健的机器人错误管理系统提供见解。我们的研究结果有助于理解人机交互中的发展差异，并为设计能够针对多种失败类型有效进行错误恢复的儿童友好型机器人系统提供参考。

2 相关工作

我们总结了相关研究，探讨如何为日益多样化的用户群体制定有效的机器人错误检测与恢复策略。

机器人错误。 随着与机器人互动的机会不断增加，发生机器人错误的可能性也随之上升。人机交互领域中的机器人故障正成为一个日益兴起的研究方向 [6, 17, 39]。人机交互文献通常将“性能错误”与“社交错误”区分开来：前者会削弱用户对机器人智能水平及其执行任务能力的感知（例如未能识别指令），而后者则会违反社会规范，并削弱用户对机器人社会情感能力的感知（例如不恰当地打断用户） [44]。

研究记录了人类在多种模态下对机器人故障的不同反应 [9, 36, 49]。这些反应包括言语调整，例如重新组织提示语 [27, 39]„ 以及通过面部表情[2, 7, 18, 21, 38], 、注视模式的变化 [8], 和肢体语言调整 [15, 21, 45]表现出的情绪表现。错误的严重程度和类型会显著影响这些反应模式 [47], ，并最终影响用户信任和继续交互的意愿 [12]。

具体而言，关于连续发生的错误，Liu 等人 [27]发现，成年人会通过重新提示和语音调整来适应沟通，情绪也会从困惑演变为沮丧，并且在反复失败后可能最终放弃互动。尽管已有研究探讨儿童遇到类似情境时的反应，但结果似乎多样且复杂，仍有必要进一步研究。

儿童与机器人互动中的错误 。 儿童并非只是缩小版的成年人；他们持续发展的神经、生理和认知能力，使其与机器人的互动呈现出不同的特点 [32]。研究表明，儿童比成年人更容易将机器人拟人化 [4] ，并且可能对机器人的局限性表现出更高的容忍度。这种容忍度有多种表现：Lemaignan 等人 [25] 发现，相较于行为可预测的机器人，儿童对行为异常的机器人更感兴趣；而 Yadollahi 等人 [50] 指出，机器人的错误可能会促使儿童纠正不准确之处，从而提高他们的参与度。

儿童在与机器人交流时表现出非凡的坚持不懈，会想办法解释并为机器人的故障开脱 [46]。此外，儿童可能较难察觉某些错误类型 [11], ，而工程师显而易见的技术局限性，在儿童相信机器人具备类人能力时可能不会被发现 [28]。然而，情况并非一片乐观：一些研究表明，反复出现的机器人故障会加剧沮丧 [43] ，或削弱信任 [14]。这些相互矛盾的证据凸显了系统性研究儿童如何应对不同错误类型的连续机器人故障的必要性。

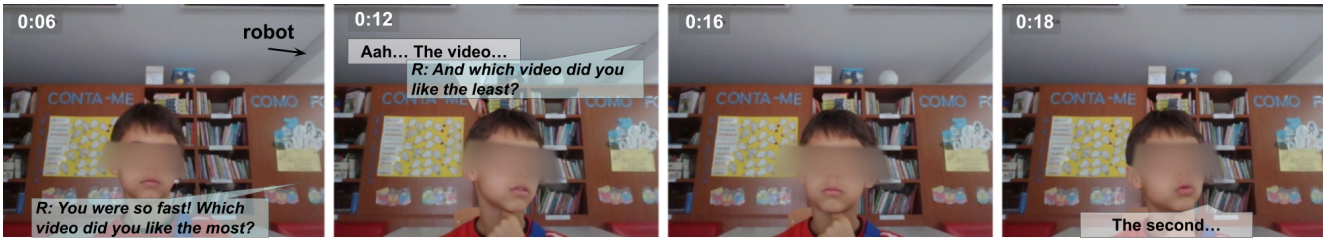

图3：对社交错误（打断）的反应。当儿童开始回答机器人提出的第一个问题时，机器人打断了他。儿童没有表现出任何外在迹象，表明他意识到这是一种社交错误行为。相反，他短暂思考了一下，然后回答第二个问题。完整的社交错误互动可通过视频图查看。

3 用户研究

为研究对话失败对儿童的影响，我们采用了 Liu 等人 [27]提出的调整版实验方案。我们的研究探讨了失败的两个维度：打断（社交错误）和未能理解（性能错误）。对于社交错误，我们测试了两种对话情境：在打断情境中，机器人（由向导操控）在参与者回答完先前的问题之前就继续进行。在对照情境中，机器人不会打断儿童。此外，我们还研究了儿童在连续遭遇机器人性能错误（未能理解）时的行为模式。

3.1 研究流程

研究设置和各个阶段如图2所示。参与者通过当地学校网络招募。所有儿童均持有由其法定监护人签署的书面同意书，并被告知可以随时退出活动。最初，参与者进入一间独立的教室，在那里认识了“机器人 Simon”（Nodbot）。他们被要求坐在椅子上，面向一台笔记本电脑，并按照问卷中的说明以及机器人提供的指示进行操作。与机器人进行简短交谈（“你好，你叫什么名字？”、“你以前和机器人交谈过吗？”）后，参与者填写人口统计信息，随后观看一组精心平衡的6段短视频，视频中包含机器人和人类，且结果有成功也有失败（完整列表见补充材料）。之后，参与者填写一份关于机器人感知的小问卷（更多详情见第3.3小节）。接着，参与者会大声告诉 Simon 自己已经完成。机器人会向他们提出一系列问题（“你最喜欢哪个视频？”、“最不喜欢哪个？”），期间机器人会表现出打断 或对照 行为。之后，机器人会指示参与者请它（机器人）把研究人员叫回房间（“我们完成了！ 先不要起身，告诉我去叫研究人员。”）。这台由向导操控的机器人经过编程，会导致互动失败。参与者与 Simon 之间的互动预期按以下方式进行：

(1) 参与者请 Simon 呼叫研究人员。(2) Simon 回复 “抱歉，我不明白”（错误 I）(3) 参与者再次提醒 Simon。(4) Simon 回复 “抱歉，我不明白”（错误 II）(5) 参与者再次提醒 Simon。(6) Simon 回复 “抱歉，我不明白”（错误 III）(7) 参与者再次提醒 Simon。(8) Simon 回答“好的，我会呼叫研究人员”。

最后，研究人员进入了房间。研究人员不在房间内时，要求儿童再次填写问卷，以评估互动对机器人感知的影响。最后，参与者接受了事后说明，并获悉了此次欺骗。未提供补偿。完成整个流程大约需要 7 分钟。这些数据是在康奈尔大学机构审查委员会豁免协议 #IRB0010006 下收集的。该研究流程已获里斯本大学伦理委员会批准。

3.2 系统实现

“Simon”机器人命名为Nodbot[1]，使用伺服电机沿两个轴（纵向和横向）进行控制，并可通过蓝牙扬声器发出声音。房间内的互动由两台摄像机记录：一台拍摄房间全景，另一台为笔记本电脑摄像头。互动由向导操控。参与者在另一台笔记本电脑上填写问卷并观看视频。

3.3 测量指标

为评估儿童对机器人的感知，我们发放了一份简短问卷，其中包含五个问题，用于测量人机交互的关键维度。所有问题均采用 5 点李克特量表（1 = 完全没有， 5 = 非常多），并以葡萄牙语呈现。表 1 列出了完整的问题及其对应的测量维度。这些问题基于此前有关测量机器人感知的文献 [3, 19, 29] ，针对的维度包括继续互动意愿、能力、信任、社会接受度和喜爱度。这些问题的克朗巴赫α系数为 0.8。

我们计算了每位参与者的差异分数（互动后减去互动前），并使用独立样本检验比较了两种实验条件下的变化。根据试点测试的结果，我们在连续失败序列之后，而非紧接着中断行为之后，发放了互动后问卷，因为简短的互动时长（三个问题）会过于频繁地打断实验流程，而且初步结果显示这两个测量时间点之间的差异很小。这种方法能够捕捉中断行为与连续性能错误对机器人感知产生的叠加影响。

[1]基于 Rei Lee 的设计：https://infosci.cornell.edu/~reilee/NodBot/



图 4：面对连续性能错误时的回应。儿童通过在更长的提示中靠近机器人、表现出沮丧来回应机器人的错误；第三次错误后，还会寻找研究人员。经过较长的响应时间后，儿童终于重复了指令，转而使用礼貌用语并微笑以示幽默。

表 1：机器人感知问卷条目及测量维度。

问题	维度
从 1 到 5，你有多想要再次交谈 机器人 Simon?	继续互-动的意愿
从 1 到 5，你认为机器人 Simon 知道如何与你交谈吗?	能力
从 1 到 5，你有多信任机器人 Simon?	信任
从 1 到 5，你有多想成为朋友与机器人 Simon?	社会接受度
从 1 到 5 分，您觉得机器人有多令人愉悦 Simon?	可喜爱性

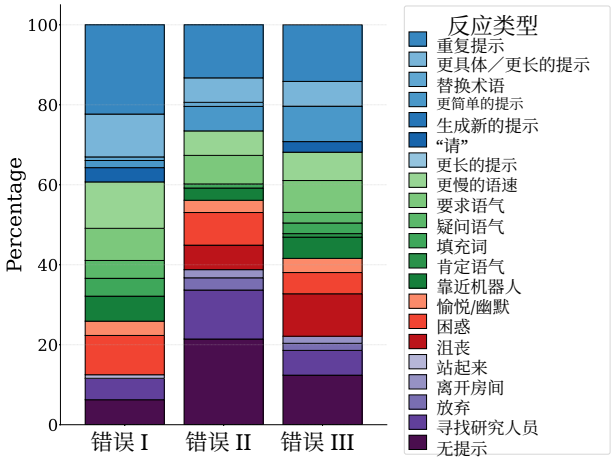

图 5：儿童对连续机器人错误的行为反应分布。展示了不同反应类型的相对频率。

3.4 反应标注

研究人员分析了每位参与者的视频（笔记本电脑摄像头拍摄），以识别常见的反应模式。使用 ELAN [2] 进行标注。对于**社交错误**（打断），我们标注了以下情况：a）儿童继续回答第一个问题（在打断前提出）；b）儿童回答新问题；c）儿童不再参与互动。

对于**性能错误**（连续未能理解），我们改编了 [27], 中的代码，这些代码记录在代码手册中，可作为补充材料获取。代码包括 **口头重新提示**（例如，在提示语中使用更具体的词语、使用更简单的提示语）、**调整语气或语速节奏**（例如，说话更慢、使用要求更高的语气）、**情绪表现**（例如，困惑、沮丧）以及**脱离参与**互动（例如，站起来、寻找研究人员）。

[2]hhttps://archive.mpi.nl/tla/elan

3.5 参与者

该研究在社交错误的两种实验条件下共纳入59名参与者：其中30名处于打断条件，29名处于控制条件。所有参与者均经历了连续的性能错误行为。参与者年龄为8至10岁（8岁：n=35，9岁：n=21，10岁：n=3），其中男孩32名，女孩27名。几乎所有参与者都是葡萄牙籍人士，只有3个例外：1名乌克兰人、1名美国人和1名巴西人。由于技术问题或背景噪声，我们将6名参与者排除在社交错误分析之外，将9名参与者排除在性能错误反应分析之外。

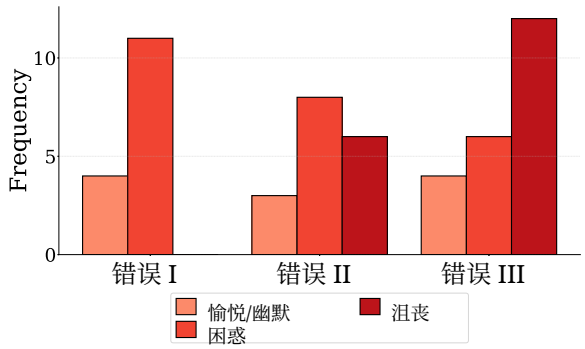

图6：连续机器人错误中情绪反应的频率。困惑在错误 II之后达到峰值，而沮丧逐渐增加，并在错误 III时占据主导。愉悦/幽默在所有错误中均保持相对一致。

4 结果

下面，我们展示社交错误和性能错误如何影响对机器人的感知，并对连续的机器人性能错误开展广泛且比较性的行为分析。我们还提供了一段补充视频作为附加材料。

4.1 机器人错误对感知的影响

对于社交错误互动，在控制条件下，3名儿童未与机器人互动，其余25名参与者自然地回答了机器人的问题。在打断条件下，21名儿童（84%）转而回答新提出的问题，没有承认打断的发生（示例见图3），仅有2名儿童继续回答之前的问题。另有2名儿童未与机器人互动。

我们通过评估各条件下的评分差异，检验打断是否影响机器人感知。我们使用Shapiro-Wilk检验评估差异分数分布的正态性。由于所有问卷条目均不满足正态性假设（所有 $p < 0.05$），因此我们采用非参数Mann-Whitney U检验比较打断条件与控制条件之间的差异分数。在对多重比较应用Bonferroni校正后（$\alpha = 0.01$），条件之间的机器人感知不存在统计学差异（完整结果见补充材料）。鉴于实验条件之间不存在显著差异，后续分析基于合并数据集进行，将所有参与者视为同一组，而不考虑其被分配到的条件。

为考察机器人互动本身是否会影响儿童的感知，而不论实验条件如何，我们对所有参与者在互动前后的反应进行了配对比较（$n = 52$）。Shapiro-Wilk检验表明，所有差异分数均不符合正态分布（均为 $p < 0.05$），因此我们针对每个问卷条目采用了Wilcoxon符号秩检验。在应用Bonferroni校正进行多重比较后，没有任何比较接近统计显著性（$\alpha = 0.01$），这表明错误的机器人互动

在所有测量维度上，都没有对儿童对机器人的看法产生有意义的影响。

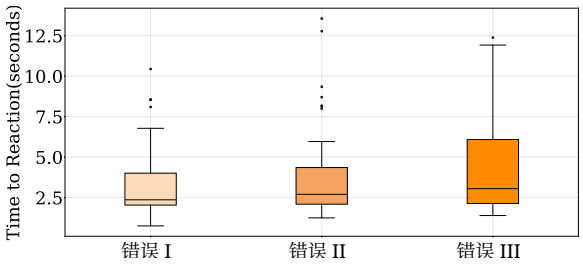

图 7：连续机器人错误下的响应延迟。随着错误连续发生，响应时间延长且变异性增大，其中错误 III 的延迟最长、变异性最大。

4.2 对连续（性能）错误的反应

我们分析了视频，并记录了儿童对连续性能错误（未能理解）的回应中呈现的行为模式，以下将对这些模式进行说明，并通过案例片段和一段补充视频图示加以展示。我们还将这些反应模式与 Liu 等人报告的模式进行比较，以推断这些模式的普适性。 [27],

与机器人连续错误互动的平均时长为 36.28 ± 16.07（$M \pm SD$）秒。分析共使用了 402 条标注（详情见补充材料）。

4.2.1 儿童行为反应。 我们发现，儿童面对连续机器人故障时，在言语、情绪和参与度方面呈现出不同的反应模式。图 5 展示了连续错误中各类反应的分布。

言语交流策略。 儿童采用了各种重新提示策略，并调整提示内容和表达方式。他们会修改提示语（重复、简化或增加具体信息），改变声音语调（放慢语速，采用命令式或疑问式语气），有时还会从 assertive 命令（“呼叫研究人员。”）转为礼貌请求（“您能帮忙呼叫研究人员吗？”）（图 1、图 4）。

情绪反应演变。 情绪表现随着错误的累积显著变化。错误 I 后，困惑占主导；错误 II 后，困惑仍然存在；到错误 III 时，沮丧达到峰值（图 6）。所有错误中都持续出现 amusement，但总体比例相对较低。

脱离行为。 一项关键发现是，儿童逐渐脱离机器人互动。错误 II 后，脱离行为显著增加，其中“无提示”回应和“寻找研究人员”最为常见（图9）。一些儿童完全停止互动，导致实验提前终止，未能出现全部三个错误（图10）。

参与模式动态. 为了解儿童参与动态，我们考察了他们在连续发生的每次错误中的参与—脱离模式。我们将参与定义为与机器人进行任何言语互动（包括语调/节奏变化或靠近机器人），而脱离参与 则包括注意力从机器人身上转移的行为（站起来、寻找研究人员、离开房间）。大多数儿童最初都参与了互动，但

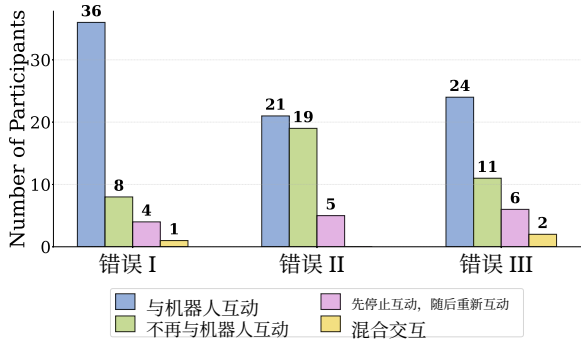

图 8：儿童在连续机器人错误中的参与模式。蓝色柱表示持续与机器人互动，绿色柱表示脱离机器人互动，粉色和黄色表示复杂参与模式：也就是说，对于单次错误，儿童可能最初表现出脱离参与（例如寻找研究人员），随后又与机器人互动（例如与机器人交谈）；混合 参与意味着儿童切换行为类型超过两次。从错误 I 中以参与行为为主，显著转变为错误 II 和 III 中的混合模式。

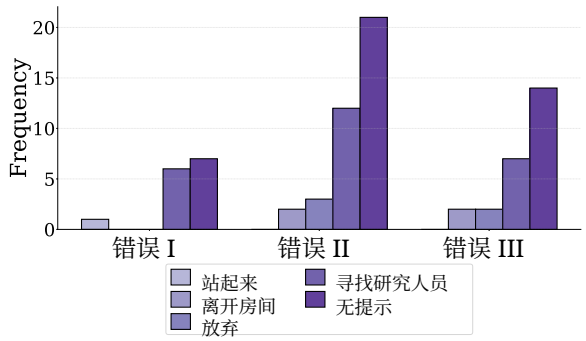

图9：连续机器人错误中脱离交互行为的频率。图示脱离策略逐步增加；到错误 III 时，“无提示”和“寻找研究人员”最为常见。其他行为包括站起来、离开房间以及完全放弃互动。

在不同错误中表现出多种模式（图 8）。最常见的三种参与轨迹（分别指错误 I、错误 II 和错误 III 时的行为）是：

- 参与 → 参与 → 参与（$n = 15$）
- 参与 → 脱离参与 → 参与（$n = 5$）
- 参与 → 不参与 → 不参与（$n = 4$）

响应时间变化。 最后，我们考察了响应时序——也就是每次机器人出错（“抱歉，我不明白”）后，儿童需要多长时间才会做出响应。图 7 展示了我们的分析结果。随着错误接连发生，响应延迟逐渐增加；错误 III 的响应时间最长且变化最大，这表明随着失败不断累积，儿童需要更多处理时间。

4.2.2 儿童连续性能错误与成 人。

与成人行为的相似之处。 若干响应模式与先前关于成年人的研究结果一致 [27]:

- **言语适应**：儿童和成年人在遇到失败时都会修改提示语，并调整声音语调和语速节奏。
- **情绪演变**：在连续错误中，情绪都呈现出从困惑到沮丧的相似变化。
- **响应时序**：随着连续失败，两组的响应延迟和变异性均有所增加。
- **交互放弃**：两项研究中均有部分参与者在第三次错误后完全停止互动。

与成人行为的关键差异。 儿童表现出若干独特的行为模式：

- **更明显的脱离参与**：儿童表现出显著更多的脱离行为，常会暂时不理会机器人。
- **寻求外部帮助**：与成年人不同，儿童会寻找或呼叫研究人员，表明他们对出现问题的互动有着不同的能动性预期。
- **礼貌策略**：初始指令失败后，儿童更容易转而使用礼貌用语（“请”）。
- **更早脱离参与**：成年人通常主要在错误 III 后（n=7）放弃互动；而儿童从错误 I 开始便表现出脱离参与。
- **混合参与模式**：儿童表现出更为动态的参与模式，有时会先脱离参与再重新参与，这不同于成年人所表现出的更线性的进程。

这些差异表明，与成年用户相比，儿童在机器人故障恢复方面对权威、寻求帮助和社会交互规范有着不同的预期。

5 讨论

本研究的独特贡献在于，系统地将 Liu 等人的连续错误范式应用于儿童，从而能够在受控条件下直接比较成人与儿童的行为。现有研究考察了自然情境下的长期错误反应，而我们的简短、受控实验方案则隔离了对连续失败的即时反应，并揭示了儿童特有的行为，这些发现对错误恢复设计具有重要启示。本研究通过提供早期阶段反应的聚焦快照，弥补了纵向研究的不足，并连接了成人与儿童的人机交互研究。根据我们的观察得出的设计启示总结于表2中。

机器人错误未改变儿童对机器人的感知。我们的问卷调查结果表明，尽管经历了社交错误（打断）和表现失败（反复误解），儿童对机器人的整体看法仍保持稳定。这种稳健性与先前研究相符：研究指出，儿童对机器人的对话预期更加灵活 [32], ，且参与者大多没有注意到或忽略了打断行为，



图10：缺乏言语参与。在这段情境中，儿童无视机器人的指令，立即 寻找研究人员。在反复出错后，儿童表现出困惑 (0:14)、amusement (0:22)，并继续回头看或看着屏幕。

这一发现也与更广泛的证据一致，即机器人错误不一定会削弱儿童对机器人的看法 [41, 51]。Stower 等人 [42] 发现，儿童对机器人可靠性的感知与他们将机器人视为社会互动伙伴的感知彼此独立；而 Ligthart 等人[26]发现，据称只有20%的儿童会因机器人错误而感到“恼火”，其余儿童则接受了这些错误，认为机器人“还在学习”。这些研究表明，与成年人相比，儿童在面对技术局限性时，或许具有维持积极机器人关系的发展优势。不过值得注意的是，Ligthart 等人 [26] 也报告称，识别错误与舒适度/社会吸引力之间存在负相关，这表明在更长的互动情境中可能会产生关系层面的影响。

儿童对机器人故障的独特反应模式。 儿童遭遇连续机器人故障时表现出鲜明的行为模式，既有普遍的人类反应，也采用了符合其发展阶段的特定策略。从困惑到沮丧的过程与成人反应相似 [27], , 但儿童展现出独特的应对机制。他们采用的各种重新提示策略——重复、简化或改变声音语调——并非只是对程序错误的反应，也体现了会话修复 [46]。这些行为表明，即使机器人看起来无法胜任会话，儿童仍会积极运用学到的社会图式与机器人互动，尝试恢复会话连贯性。初始指令未能奏效时，儿童会使用礼貌用语形式（“请”），这表明他们将人际交往中的社会礼貌规范迁移到了与机器人的互动中，反映出他们对社会等级及适当沟通策略的理解正在发展 [13]。

寻求支援。 一个关键的区别特征是，儿童经常向外寻求帮助，尤其是寻找或呼叫研究人员。这种“寻求支援”模式与先前在机器人辅助语言学习场景中的报告相符 [48]，且与成年参与者的行为明显不同他们主要尝试自主解决问题 [27],，这可能反映了感知能动性和权威关系方面的发展差异：儿童期望成年人为复杂问题提供解决方案。Serholt 等人 [34] 也记录了儿童对机器人错误的反应，包括靠近机器人、进行言语适应，以及尤为重要的脱离参与和求助行为。这些行为并不代表互动失败，而是体现了儿童的适应性问题解决能力，并对教育或治疗情境具有重要意义；在这些情境中，求助行为可被视为积极参与的指标。

逐步脱离与重新参与的动态过程。儿童的参与模式比成年人的线性变化过程更为动态，参与者会在参与和脱离参与状态之间来回切换。Oertel 等人 [30] 报告称，随着新奇效应减弱，儿童可能会减少与机器人的互动；但我们的互动时间较短，且观察到了重新参与的模式，这表明其中可能存在不同的机制。相反，我们将这些脱离参与的模式解读为主动的问题解决和适应。当工具无法按预期发挥作用时，有能力的使用者可能会弃用它，并寻找替代策略来实现目标。在本例中，儿童的首要目标是让研究人员重新进入房间。机器人不过是实现这一目标的工具。当机器人反复出错时，儿童理性地不再使用这个有故障的工具，转而采取更有效的策略：向人类寻求帮助。这种解读将行为从“机器人故障”重新界定为“使用者能力”，展现了儿童精细的元认知能力：他们能够评估工具局限性，并调整问题解决方式。相较之下，成年人可能因为对互动情境属于一项科学研究有着更成熟的理解，而不愿表现出这种亲自动手解决问题的行为。参与与脱离参与之间的灵活转换表明，儿童能够从容应对机器人故障，即使反复遭遇错误，仍愿意继续互动 [33]。

表2：基于本研究及既有研究得出的儿童—机器人交互中有关机器人错误的设计启示 [26, 27, 33, 41, 48]。

观察到的儿童反应／洞见	针对儿童—机器人交互的直接设计启示
尽管机器人接连出错，儿童对机器人的整体感知仍然保持稳定。	**容错设计（感知韧性）：**重点在于妥善处理错误，而不是积极地消除这些错误。允许发生轻微、非关键的错误，而不立即进行冗长的道歉，借助儿童灵活的预期。
儿童经常会寻找或呼叫研究人员／成人来帮忙，以应对反复失败（呼叫支援）。	**促进寻求外部帮助（能动性转移）：**识别明确寻求人类协助的请求（言语或非言语）作为有效的适应性策略，并实现平稳的机制以移交控制权或召唤人类帮助。
儿童会运用复杂的言语策略：重复、逐步提高具体程度，最终采用礼貌请求（“请”）。	**实施多层次对话修复：**设计高度完善的听觉输入系统对提示改述（语速节奏、语气、语言复杂度的变化）较为敏感，并提高错误基于礼貌标记（例如“请”）的置信度。
儿童表现出动态的参与模式，在脱离参与与重新参与机器人之间循环。	**采用非侵入式错误恢复（容忍停顿）：**将脱离参与视为一种主动问题解决阶段。避免立即采用具有侵扰性的重新参与协议；而应被动监测并允许儿童自然地重新参与。
随着错误接连发生，响应延迟会增加。	**利用时序作为失败指标：**将响应延迟分析（机器人错误发生与儿童后续响应之间经过的时间）作为实时、客观的指标，以衡量认知负荷和痛苦的加剧。并将儿童后续响应作为实时、客观的指标，以衡量认知负荷和痛苦的加剧。

方法论局限与情境约束。这种高度受控的互动环境采用了一种特定的对话失败范式，仅代表机器人错误情境的一种类型；尽管其他文献也支持我们的发现，但这些发现很大程度上受实验环境和儿童发展阶段的具体情境所限。此外，在公立学校收集数据需要对实验进行大量调整，也降低了对环境的控制，包括意料之外的打断，导致部分参与者被排除。儿童对实验指令的遵从程度各不相同，有些儿童选择完全忽略机器人的互动邀请。这反映了儿科研究中真实存在的挑战，却也使系统性分析更加复杂。参与者的年龄范围有限且文化同质性较高，进一步限制了研究结果的可迁移性 [32, 37]。最重要的是，在连续出现错误期间，互动时间较短（平均 36.28 秒），只能呈现儿童与机器人的互动动态的一个片段；延长互动时间或许会带来截然不同的适应模式。

伦理考量与数据保护的复杂性。该研究凸显了涉及儿童数据的国际协作研究所面临的重大挑战，尤其是在数据收集受欧洲数据保护框架约束的情况下。儿童数据保护方面的严格要求虽然对参与者福祉绝对必要，却也为数据共享和协作分析设置了重大障碍。这些保护措施确保了对弱势参与者进行适当的保护，但同时也限制了跨文化验证和更广泛协作洞见的机会。

本研究还通过绿野仙踪式控制采用了轻度欺骗：儿童相信机器人是自主运行的，而研究人员则控制其故障。虽然这种欺骗有时对于引发自然反应是必要的，并且在得到妥善管理时符合伦理要求，但对儿童参与者使用时仍需谨慎考量。所有儿童都接受了充分的事后说明，研究人员向他们解释了机器人的真实运行方式和研究目的，确保他们理解并未发生实际故障，并明白他们的反应对于改进机器人设计而言是有价值的贡献。未来涉及儿童欺骗的研究应继续优先采用透明的事后说明程序，并确保暂时性的欺骗不会使他们对技术能力产生持久的困惑或痛苦。

未来方向与更广泛的应用。 未来研究应探讨儿童错误容忍度和恢复策略的纵向发展，考察文化差异对机器人错误反应的影响，并研究不同错误类型（社交错误和性能错误），以制定更具包容性的设计实践。观察到的求助行为值得进一步研究，以判断其是否可作为人机协作学习环境中的潜在积极指标。

尝试通过儿童的行为信号自动进行错误检测时，出现了一个重大的实际挑战。与保持坐姿并处于摄像机画面内的成年人不同，儿童经常进行不可预测的移动，包括站立、转身或以意料之外的角度调整位置。这种动态行为导致计算机视觉和姿态估计算法产生大量缺失数据，使传统的面部表情和身体姿态分析变得不可靠。未来用于儿童—机器人交互的自动化系统将需要专门的方法，包括采用多个摄像机角度、能够处理遮挡和快速移动的稳健跟踪算法，以及可能超越视觉分析的替代传感模态。这些限制凸显了针对儿童的专门行为信号处理的必要性，即需要考虑儿童天生更加动态且不可预测的身体行为。

6 结论

本研究考察了儿童如何应对连续机器人故障。观察到的求助行为和感知韧性表明，与此前的认识相比，儿童可能是人机协作中适应能力更强的伙伴，尽管这也为自动错误检测系统带来了独特的技术挑战。随着机器人日益进入涉及儿童的社会情境，这些洞见对于创建有效、可持续且符合发展阶段的人机交互系统至关重要。

致谢

本项目经 Lasige UID/00408/2025 和 FCT/PEX 2024.15729.PEX 资助。感谢参与数据收集工作的学校。

参考文献

[1] Deepali Aneja、Daniel McDuff 和 Mary Czerwinski。2020。《人机交互中的会话错误分析》。载于第20届ACM智能虚拟代理国际会议论文集（线上活动，苏格兰，英国）（IVA '20）。计算机协会，纽约，纽约州，美国，文章 3, 8 10 1145 3383652 3423901 页。doi: . / 。Reuben M. Aronson。。《用于人机共享操作中错误检测的凝视》。载于 RSS Workshop: Towards a Framework for Joint Action。[3] Christoph Bartneck、Dana Kulic、Elizabeth Croft 和 Susana Zoghbi。2008。《机器人拟人化、生命性、可喜爱性、感知智能和感知安全性的测量工具》。《国际社交机器人学杂志》 1 01 2008 71 81 10 1007 12369 008 0001 3 （），–。doi: . /s - - - Tony [4] Belpaeme、Paul Baxter、Joachim Greeff、James Kennedy、Robin Read、Rosemarijn Looije、Mark Neerincx、Ilaria Baroni 和 MattiaCoti Zelati。2013。《儿童与机器人的互动：视角与挑战》。452–459。doi:10.1007/978-3-319-02675-6 45.[5]Tony Belpaeme、James Kennedy、Aditi Ramachandran、Brian Scassellati 和 Fumihide Tanaka。2018。《用于教育的社交机器人：综述》。《科学·机器人学》3，21（2018），eaat5954。arXiv:https://www.science.org/doi/pdf/10.1126/scirobotics.aat5954 doi:10.1126/scirobotics.aat5954[6] Alexandra Bremers、Alexandria Pabst、Maria Teresa Parreira 和 Wendy Ju。2024。《运用社交线索识别人机交互中的任务失败：概述、最新进展与未来方向》。arXiv:2301.11972 [cs.RO] https://arxiv.org/abs/2301.11972[7] Alexandra Bremers、Maria Teresa Parreira、Xuanyu Fang、Natalie Friedman、Adolfo Ramirez-Aristizabal、Alexandria Pabst、Mirjana Spasojevic、Michael Kuniavsky 和 Wendy Ju。2023。《用于人机交互中失败检测的旁观者情感检测（BAD）数据集》。载于2023年IEEE/RSJ智能机器人与系统国际会议（IROS）。IEEE，11443–11450。doi:10.1109/IROS55552.2023.10342442[8] Dito Eka Cahya、Rahul Ramakrishnan 和 Manuel Giuliani。2019。《人机协作中无错误与错误情境下社交信号的静态与时间差异》。载于《社会机器人学：第11届国际会议，ICSR 2019，马德里，西班牙，2019年11月26–29日，论文集》（马德里，西班牙）。Springer-Verlag，柏林，海德堡， 189 199 10 1007 978 3 030 35888 4_18 –。doi: . / - - - - [9] Kate Candon、Nicholas C. Georgiou、Helen Zhou、Sidney Richardson、Qiping Zhang、Brian Scassellati 和 Marynel Vázquez。2024。《REACT：用于分析人类对机器人反应及评价性反馈随时间变化的两个数据集》。载于2024年ACM/IEEE人机交互国际会议论文集（博尔德，科罗拉多州，美国）（HRI '24）。计算机协会，纽约，纽约州，美国， 885 889 10 1145 3610977 3637480 –。doi: . / 。Huili Chen、Hae [10] Won Park 和 Cynthia Breazeal。。《与儿童共同教学和学习：互惠同伴学习与社交机器人对儿童学习和情感投入的影响》。《计算机与教育》150（2020），103836。doi: 10 1016 2020 103836 . /j.compedu. 。Eve V. Clark。。《第一语言习得》（第2版）。剑桥大学出版社。[12] Connor Esterwood 和 Lionel P. Robert Jr. 2023。《三次机会，用完出局！多次人机信任违背与修复对机器人可信度的影响》。《计算机与人类行为》142（2023），107658。doi:10.1016/j. chb.2023.107658[13] Catherine Garvey 和 Ginger Berninger。1981。《儿童对话中的时机与轮流交谈1》。《话语过程》4，1（1981），27–57。arXiv:https://doi.org/ 10 1080 01638538109544505 10 1080 01638538109544505 . / doi: . / [14] Denise Y. Geiskkovitch、Raquel Thiessen、James E. Young 和 Melanie R. Glenwright。2019。《什么？那不是椅子！机器人信息错误如何影响儿童对机器人的信任》。载于2019年第14届ACM/IEEE人机交互国际会议（HRI）。48–56。doi:10.1109/HRI.2019.8673024[15] Manuel Giuliani、Nicole Mirnig、Gerald Stollnberger、Susanne Stadler、Roland Buchner 和 Manfred Tscheligi。2015。《对不同人机交互研究视频数据的系统分析：错误情境下社交信号的分类》。《心理学前沿》6（2015），931。[16] Goren Gordon、Samuel Spaulding、Jacqueline Kory Westlund、Jin Joo Lee、Luke Plummer、Marayna Martinez、Madhurima Das 和 Cynthia Breazeal。2016。《面向儿童第二语言技能的社交机器人导师的情感个性化》。AAAI人工智能大会论文集30，1（3月 2016 10 1609 30 1 9914 ）。doi: . /aaai.v i 。Shanee Honig 和 Tal Oron-Gilad。[17] 2018。《理解并解决人机交互中的失败：文献综述与模型构建》。《心理学前沿》第 9 2018 2018 10 3389 2018 00861 卷-（）。doi: . /fpsyg. 。[18] Kao-Shing Hwang、JL Ling、Yu-Ying Chen 和 Wei-Han Wang。2014。《通过情绪表达为强化学习进行奖励塑形》。载于2014年IEEE系统、人类与控制论国际会议（SMC）。IEEE，IEEE，纽约，纽约州，美国，1288–1293。[19] C.D. Kidd 和 C. Breazeal。2004。《机器人对用户感知的影响》。载于2004年IEEE/RSJ智能机器人与系统国际会议（IROS）（IEEE目录号04CH37566），第4卷。3559–3564，第4卷。doi:10.1109/IROS.2004.1389967

[20] Dimosthenis Kontogiorgos、Andre Pereira、Boran Sahindal、Sanne van Waveren 和 Joakim Gustafson。2020。对机器人对话失败的行为反应。载于《2020年第15届ACM/IEEE人机交互国际会议（HRI）》。IEEE，IEEE，美国纽约州纽约市，53–62。[21] Dimosthenis Kontogiorgos、Minh Tran、Joakim Gustafson 和 Mohammad Soleymani。2021。对人类针对机器人对话失败的反应进行系统性跨语料库分析。载于《2021年多模态交互国际会议论文集》。112–120。[22] Dimosthenis Kontogiorgos、Sanne van Waveren、Olle Wallberg、Andre Pereira、Iolanda Leite 和 Joakim Gustafson。2020。与发生故障的机器人互动时的具身效应。载于《2020年CHI计算系统中的人为因素会议论文集》（美国夏威夷州檀香山）（CHI '20）。计算机协会，美国纽约州纽约市， 1 14 10 1145 3313831 3376372 –。doi: 。/。S H [23] Landry、K E Smith、P R Swank 和 C L Miller-Loncar。。早期母亲与儿童的影响对儿童日后独立认知与社会功能的作用。《儿童发展》71（2000），358–375。[24] Séverin Lemaignan、C E R Edmunds、E Senft 和 T Belpaeme。2018。PInSoRo数据集：支持对儿童间及儿童与机器人间社会动态的数据驱动研究。PLoS ONE 13, 10 2018 0205999 10 1371 0205999 ( ), e 。[25] doi: 。/journal.pone。Séverin Lemaignan、Julia Fink、Francesco Mondada 和 Pierre Dillenbourg。。你做错了！研究儿童-机器人互动中的意外行为。载于《社会机器人学》，Adriana Tapus、Elisabeth André、Jean-Claude Martin、François Ferland 和 Mehdi Ammi（编）。施普林格国际出版公司，尚，390–400。[26] Mike E. U. Ligthart、Mark A. Neerincx 和 Koen V. Hindriks。2022。相互熟悉：儿童-机器人关系形成的初步步骤。《机器人学与人工智能前沿》第 9 2022 2022 10 3389 2022 853665 卷-（）。doi: 。/frobt. 。[27] Shannon Liu、Maria Teresa Parreira 和 Wendy Ju。。“我不干了”：描述人类对连续机器人故障的反应。载于《2025年第20届ACM/IEEE人机交互国际会议（HRI）》。1458–1462。doi:10.1109/HRI61500.2025.10974098[28] Marco Nalin、Linda Bergamini、Alessio Giusti、Ilaria Baroni 和 Alberto Sanna。2011。儿童在轻度受限环境中对机器人伴侣的感知。[29] Isabel Neto、Yuhan Hu、Filipa Correia、Filipa Rocha、João Nogueira、Katharina Buckmayer、Guy Hoffman、Hugo Nicolau 和 Ana Paiva。2024。“我才没碰你呢，是机器人！”：借助触觉机器人促进不同视觉能力儿童的包容。载于《2024年ACM/IEEE人机交互国际会议论文集》（美国科罗拉多州博尔德）（HRI '24）。计算机协会，美国纽约州纽约市，511–521。doi:10.1145/3610977.3634992[30] Catharine Oertel、Ginevra Castellano、Mohamed Chetouani、Jauwairia Nasir、Mohammad Obaid、Catherine Pelachaud 和 Christopher Peters。2020。人机交互中的参与度：概述。《机器人学与人工智能前沿》第7卷-2020（2020）。doi:10.3389/frobt.2020.00092[31] Raquel Ros、Elettra Oleari、Clara Pozzi、Francesca Sacchitelli、Daniele Baranzini、Anahita Bagherzadhalimi、Alberto Sanna 和 Yiannis Demiris。2016。以动机为导向的方法：支持长期儿童-机器人互动中的健康习惯。《国际社交机器人学杂志》8，5（2016），599–617。doi:10.1007/s12369-016-0356-9[32] Irina Rudenko、Andrey Rudenko、Achim J Lilienthal、Kai O Arras 和 Barbara Bruno。2024。儿童因素在儿童-机器人交互中的作用：探究发展阶段和个体特征的影响。《国际社交机器人学杂志》16, 8 2024 1879 1900 10 1007 12369 024 01121 5 ( ), –。doi: 。/ [33] s - - - Sofia Serholt。。儿童与机器人导师互动中的交互故障：一项纵向研究。《计算机与人类行为》81（2018），250–264。doi:10. 1016 2017 12 030 /j.chb. 。[34] Sofia Serholt、Lena Pareto、Sara Ekström 和 Sara Ljungblad。2020。儿童-机器人交互中的故障与修复：研究小学课堂中儿童与机器人学生的复杂互动。《机器人学与人工智能前沿》第 7 2020 2020 10 3389 2020 00046 卷-（）。doi: 。/frobt. 。[35] Gabriel Skantze。。预测和调节人机对话中的参与平等：年龄和性别的影响。载于《2017年ACM/IEEE人机交互国际会议论文集》（奥地利维也纳）（HRI '17）。计算机协会，美国纽约州纽约市，196–204。doi:10.1145/2909824.3020210[36] Micol Spitale、Maria Teresa Parreira、Maia Stiber、Minja Axelsson、Neval Kara、Garima Kankariya、Chien-Ming Huang、Malte Jung、Wendy Ju 和 Hatice Gunes。2024。ERR@HRI 2024挑战赛：对人机交互中的错误与故障进行多模态检测。arXiv:2407.06094 [cs.RO] https://arxiv.org/abs/2407. 06094[37] Micol Spitale、Rebecca Stower、Elmira Yadollahi、Maria Teresa Parreira、Nida Itrat Abbasi、Iolanda Leite 和 Hatice Gunes。2023。儿童-机器人交互中可复现性的系统综述。arXiv:2309.01822 [cs.HC] https: //arxiv.org/abs/ 2309 01822 。Maia Stiber 和 [38] Chien-Ming Huang。2021。并非所有错误都一样：探究人类对不同严重程度机器人错误的反应。载于《2020年多模态交互国际会议伴随出版物》（线上活动，荷兰）（ICMI '20 Companion）。计算机协会，

美国纽约州纽约市， 97 101 10 1145 3395035 3425245 –。
[39] 2025
doi: 。/。Maia Stiber、Russell Taylor 和 Chien-Ming Huang。。通过人类反应实现机器人错误感知：实现、评估与建议。arXiv:
2501 05723 [ ] 2501 05723 。cs.RO https://arxiv.org/abs/ 。
[40]
Maia Stiber、Russell H. Taylor 和 Chien-Ming Huang。2023。利用社交信号实现灵活的错误感知型人机交互。载于《*2023*年*ACM/IEEE*人机交互国际会议论文集》（瑞典斯德哥尔摩）（*HRI '23*）。计算机协会，美国纽约州纽约市，222–230。doi:10.
1145/3568162.3576990[41] Rebecca Stower、Rania Abdelghani、Marisa Tschopp、Keegan Evangelista、Mohamed Chetouani 和 Arvid Kappas。2022。探索儿童与机器人互动中机器人犯错的空间。《交互研究》23，2（2022），243–288。
doi:10.1075/is.21034.sto[42] Rebecca Stower、Arvid Kappas 和 Kristyn Sommer。
2024。机器人犯错何时才合适？在选择性信任任务中，儿童对机器人的信任胜过对人类的信任。《计算机与人类行为》
157 2024 108229 10 1016 2024 108229 ( )， 。doi: 。/
[43] 2012
j.chb. 。Fumihide Tanaka 和 Shizuko Matsuzoe。。儿童教会接受照护的机器人，以促进自身学习：在课堂中开展词汇学习实地实验。《人机交互期刊》1，1（2012年7月），78–95。doi:10.5898/JHRI.1.1. Tanaka[44] Leimin Tian 和 Sharon Oviatt。2021。人机交互中的社交错误分类体系。《人机交互期刊》10，2，文章13（2021年2月），32页。
doi:10.1145/3439720[45] Pauline Trung、Manuel Giuliani、Michael Miksch、Gerald Stollnberger、Susanne Stadler、Nicole Mirnig 和 Manfred Tscheligi。2017。头部与肩部：人机交互中的自动错误检测。《国际会议论文集》
多模态交互会议 (2017). [46] Sherry Turkle、Will Taggart、Cory D. Kidd 和 Olivia Dasté。2006。与儿童和老年人的关系型人工制品：网络陪伴的复杂性。连接科学 18, 4 2006 347 361 10 1080 09540090600868912 ( )， –。
[47] 2019
doi:。/ Sanne van Waveren、Elizabeth J. Carter 和 Iolanda Leite。。团队合作：社交机器人协作任务中错误严重程度的影响。载于第*19*届*ACM*智能虚拟代理国际会议论文集 (法国巴黎) *(IVA '19)*。计算机协会，纽约，纽约州，美国，
151 158 10 1145 3308532 3329475 –。doi:。/。Outi Veivo 和 Maarit
[48] 2022
Mutta。。机器人辅助的第二语言学习中的对话中断。《计算机辅助语言学习》38, 1-2 (2022), 30–51。doi:10.1080/ 09588221.2022.2158203[49] Lennart Wachowiak、Peter Tisnikar、Andrew Coles、Gerard Canal 和 Oya Celiktutan。2024。基于用户反应检测人机交互中交互中断的时间序列分类流程。ACM。doi:10.1145/3678957.3688386[50] Elmira Yadollahi、Wafa Johal、Ana Paiva 和 Pierre Dillenbourg。2018。机器人中的指示性手势如何损害儿童与机器人的协作。载于第*17*届*ACM*交互设计与儿童会议论文集 (挪威特隆赫姆) *(IDC '18)*。计算机协会，纽约，纽约州，美国，195–206。
doi:10.1145/3202185.3202743[51] Burcu Ünlütabak 和 Duygun Erol Barkana。2025。儿童与社交机器人的信息搜索：聚焦儿童的提问行为。《儿童-计算机交互国际期刊》45 (2025), 100751。doi:10.1016/j.ijcci.2025. 100751