

## 《心理学报》审稿意见与作者回应

题目：自闭症谱系障碍儿童在多重线索下习得词语的眼动研究

作者：荆伟

### 第一轮

#### 审稿人 1 意见：

意见 1：1.“面朝枯燥物体目光注视枯燥物体（社会性线索与感知觉线索矛盾）”。此处“目光注视枯燥物体”是否应为“有趣物体”？不然，两种线索并非“矛盾”。如果按现在的说法，那么应如何理解“线索矛盾”？

回应：有趣物体颜色鲜艳形象可爱，而枯燥物体则毫无色彩形象呆板。在预实验中，当询问大班幼儿哪一个玩具更有趣，更喜爱哪一个玩具时，绝大部分大班幼儿都选择了有趣物体。眼动数据分析的结果亦表明，在未施加社会性线索的注意阶段，被试均对有趣物体给予了更多的注视时间，这就确保了实验材料的有效性。当年轻女性面朝枯燥物体目光注视枯燥物体时，如果依据感知觉线索，那么被试将会选取有趣物体作为新异词语的所指对象。如果依据他人视线，那么被试将会跟随他人视线选取枯燥物体作为新异词语的所指对象。也就是说，感知觉线索指向的是有趣物体，社会性线索指向的是枯燥物体，两种线索指向的物体不同，因此将这种实验条件定义为矛盾线索

意见 2：在（2）脸部局部注视时间分析这部分，此处“奇异值”是否应为“歧义值”？

回应：作者认为应该是“奇异值”，这里之所以要删掉 2 名 AD 儿童，由于兴趣区面积较小，2 名 AD 儿童在各个兴趣区内的注视时间非常少，是作为平均值 3 个标准差之外的数据予以剔除的。

意见 3：在 4.2 部分第 16 行中的“TD”是否应为“A D”？原句为“而 T D 儿童对三个区域的关注度相同……”。

回应：已修改，标注为黄色。

## 审稿人 2 意见：

**意见 1：** 被试年龄不匹配，差别太大？对于发展性的研究来说是有问题的。作者可以参照前两条参考文献中被试的匹配情况。

**回应：** 由于自闭症谱系障碍儿童群体内部存在较大的异质性，表现为他们虽然在社会交往、语言交流、刻板行为上的障碍是其共同特征，但在言语智力发展水平和非言语智力发展水平上存在较大的差异。仅有少数的自闭症谱系障碍儿童具有正常的言语与非言语智力发展水平，他们包括两类儿童一类是高功能的自闭症儿童另一类是阿斯伯格障碍儿童。评审老师所指的两个参考文献中的被试均是这两类儿童，他们与正常儿童在生理年龄、言语智力年龄和非言语智力年龄均能够匹配（详细信息见下表黄色部分）。

但多数自闭症谱系障碍儿童除了具有自闭症谱系障碍的核心障碍外还伴随不同程度的言语智力发展迟缓，因此，此类自闭症谱系障碍儿童的生理年龄一般远远高于与之具有相同言语智力发展水平的普通儿童，因而我们无法即控制此类自闭症谱系障碍儿童言语智力水平差异又控制其生理成熟差异。本研究与大量研究均是以此类自闭症儿童为被试（详细信息见下表蓝色部分）。由于本领域的研究是关于儿童语言习得能力，因此本研究领域内的所有研究均首先控制言语智力因素的影响作用，研究者一般将此类儿童与普通儿童进行言语智力水平的匹配，那么实验结果的差异就不能归因于言语智力水平的差异。对于此类自闭症儿童，匹配了言语智力发展水平就意味着其生理年龄将显著高于与之具有相同言语智力发展水平的普通儿童。至于非言语智力水平的差异，不同的研究采取了不同的处理方式，有些研究采用两个控制组，一个控制言语差异，一个控制非言语差异（例如，2007, Parish-Morris）；有些研究忽视非言语智力的差异，不做控制（例如，2010, Franken; 2009, Luyster）；而本研究所选取的控制组与实验组在言语智力和非言语智力上均能够匹配。此外，生理成熟水平的差异是否会影响实验结果呢？所有以此类自闭症谱系障碍儿童为被试的研究结果均显示，即便此类自闭症谱系障碍儿童具有更高的生理成熟水平以及更多的生活经验，但他们的社会认知技能障碍导致在完成实验任务时依然显著低于生理年龄远远小于他们的普通儿童，由此可见生理成熟水平上的差异未对实验结果产生影响作用。

不过，也有少数研究为了同时匹配此类自闭症谱系障碍儿童的智力发展水平和生理成熟水平，选取的控制组为同样具有智力发展迟滞的智力障碍儿童，这样实验组与控制组就能同时匹配生理成熟水平和智力发展水平（详细信息见下表红色部分）。但如此匹配又引来了新的问题，那就是研究者的目的是考察自闭症谱系障碍儿童的语言习得能力与正常的语言习得能力存在哪些差异，那么只有具有正常语言发展水平的普通儿童才掌握了正常的

语言习得技能，而智力障碍儿童同样具有语言发展障碍因而可能并未掌握正常的语言习得技能，因而选取智力障碍儿童为控制组可能不能实现研究目的。因此，在研究此类自闭症儿童时，大多是研究还是选取言语智力和非言语智力相匹配的普通儿童作为控制组。

年份	作者	是否伴随言语智力迟滞	实验组生理年龄	言语智力匹配的控制组生理年龄
1997	Baro-Cohen	是	AD 9.8years	智力障碍 9.1years
2003	Hennon	是	AD 3.0years	TD 1.0years
2005	Preissler	是	AD 7.8years	TD 1.9years
2005	Beily	是	AD 4.3years	TD 2.0years
2006	Mcduffie	是	ASD 2.9years	TD 1.3years
2007	Parish-Morri	是	AD 5.1years	TD 2.6years
2008	Tek	是	ASD 2.8years	TD 1.7years
2009	Luyster	是	ASD 2.6years	TD 1.7years
2010	Franken	是	ASD 9.4years	智力障碍 9.4years
2010	Nurby	否	ASD 7.2years	TD 7.1years
2011	Akechi	否	ASD 9.1years	TD 8.7years
2012	Akechi	否	ASD 8.8years	TD 8.6years
2012	Gliga	否	自闭症高危儿童 3.2years	TD 3.2years

意见 2：面孔中眼睛、鼻子和嘴巴的兴趣区界定与王乾东等人研究的界定是有差别的，而在后面的讨论中又提及两个研究的结果是一致的。作者对 AD 儿童获取社会性线索方式方面的研究只与王乾东等人（2013）的研究进行比较和讨论，而且由于两个研究关注的问题差别太大，作者的这样的比较意义不大，对该问题的研究结果缺乏汇聚性证据。

回应：在王乾东等人研究中，刺激材料是面孔图片，整个屏幕只呈现一个面孔，因而能够按照眼部、鼻部和嘴部的轮廓来划分兴趣区，而本研究中面孔只占整个屏幕的很小部分，如果按各部位的轮廓划分兴趣区，那么兴趣区的面积将会非常小，而被试的注视点落在如此小的面积上的概率是非常小的，那么有效值会很少，也就无法进行统计分析了。因此，只能将靠

近各部位的区域也划分到各部位的兴趣区内，将从眉毛到下颌的面孔区域平均分成三部分，这样保证了各兴趣区在眼动指标上的差异并非来自于兴趣区面积的差异。另外，评审老师在**批注 17**中也提到这个问题，就是：“两个研究的材料、兴趣区界定差别较大，得出一致的研究结果能说明什么问题。”笔者之所以要引用这个文献，是考虑到，这两个研究在两个不同的社会认知技能领域内（面孔识别和视线追随）说明了同一个问题，即社会性经验在塑造儿童成为社会性认知技能专家中的作用。笔者是在讨论中将视线追随上升到社会性认知技能这个更广泛的范畴内进行讨论时才将两个研究的结论进行比较的。评审老师认为这样不妥，笔者在讨论中不再将两个研究进行比较了，仅仅借鉴王乾东的研究利用专家新手理论解释异族效应，同样利用专家新手理论解释本研究中自闭症谱系障碍儿童在脸部局部注视模式上与普通儿童存在差异的现象。同时，又补充了一些 AD 儿童视线追随行为的研究结果作为本研究推论的证据，修改部分在 4.2 中标注为黄色。

**意见 3：AD 还是 ASD？**作者研究中的自闭症儿童的诊断标准与国外研究中 ASD 儿童的诊断标准是否一致？名称有何差别？因为作者研究的是特殊儿童的认知发展，样本特征应该交代清楚。

**回应：**笔者首先陈述自闭症谱系障碍概念的形成过程。1943 年，美国精神病医生里奥·坎纳首先使用了婴儿自闭症这一术语用来描述自己收治的 11 名具有异常行为问题的儿童。1979 年，Wing 和 Gould 提出自闭症是一种谱系障碍，即自闭症谱系障碍的概念，同时根据 ICD-9-R 和 1980 年的 DSM-III，将儿童自闭症与阿斯伯格综合症一并归入自闭症谱系障碍中。而随后在 1987 年制定的 DSM-III-R 中，待分类的自闭症谱系障碍被包含到自闭症谱系障碍范畴中。雷特氏疾患与儿童期崩解性疾患于 2000 年首次出现在 DSM-IV-TR 中。因此，至 2000 年，DSM-IV 与 ICD-10 在自闭症谱系障碍的分类及其诊断标准上已达成了初步的一致，认为自闭症谱系障碍（ASD）共同包含着五类障碍类型，分别为自闭症疾患（AD）、阿斯伯格疾患、待分类的广泛性发育障碍、雷特氏疾患和儿童期崩解症。ASD 并未取代 AD，ASD 只是包括了除典型的 AD 儿童之外自闭症症状并不明显的处于正常儿童和 AD 儿童之间灰色地带的儿童。自闭症谱系是按照异常行为的严重程度，即从轻微到严重呈现为一个聚集。自闭症谱系障碍各亚类型之间既存在的核心特征，同时也存在着一些细微的差别。同时“谱系”直接表明类型之间具有一种不间断的关系，而不是单纯的、分离的实体。同时虽然个体可以根据详细的诊断标准确定其是哪一种具体的自闭症谱系障碍类型，但是不同亚类型之间的界限并不是完全清晰的、绝对的，诊断个体是这一类障碍或是另一类经常只是一种判断。由于

自闭症被视为一种谱系障碍是自闭症研究的一项跨越性进步。因此，早期的研究多运用 AD 概念（见上表 Bailey, 2005; Baron-Cohen, Baldwin, & Crowson, 1997; Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, 1995; Preisslera & Carey, 2005; Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff, & Tager-Flusberg, 2007; Preissler, 2008 ;）而后期的研究大多数运用 ASD 概念（见上表 Akechi et al., 2011; Akechi, Kikuchi, Tojo, Osanai, & Hasegawa, 2012; Bean, 2010; Luyster & Lord, 2009; Franken, Lewis, & Malone, 2010; McDuffie, Yoder, & Stone, 2006; Norbury, Griffiths, & Nation, 2010; Swensen, Kelley, Fein, & Naigles, 2007; Tek, Jaffery, Fein, & Naigles, 2008）。

国外研究中 ASD 儿童的诊断标准均采用的是 DSM-IV 和 ICD-10。但两个标准仅能判断是或不是，不能给出一个分值。因而一部分研究还采用的 ADOS（Autism Diagnostic Observation Schedule,《孤独症诊断观察量表》）给出一个具体的分值，另一部分研究采用的是 CARS（Childhood Autism Rating Scale,《儿童自闭症评定量表》）给出分值，还有的研究两个量表均使用。本研究中的 ASD 儿童的认定和排除标准均来至儿童学籍信息，学籍信息显示这些儿童在入学之前都曾被医院诊断为 ASD 儿童。在参加本次实验之前，研究者又通过直接观察、与家长老师访谈、分析已有病历纪录等多种方式收集资料按照 DSM-IV 和 ICD-10 的标准以及 CARS 进行了进一步的汇聚式评估，从而进一步肯定了这一诊断。本研究之所以未采用 ADOS 量表，是由于我国尚未引入修订 ADOS。因此本研究采用的国内很早就引入修订的同样具有较好信效度得 CARS 量表对本研究中的自闭症儿童的症状进行评估。

意见 4: 严格来说，在方差分析中应该是偏 eta 的平方，即  $\eta_p^2 = 0.07$ 。

回应: 已全部修改了，标注为黄色。统计符号应斜体，请修改文中部分。

意见 5: 单单从行为指标的结果是无法做出这样的推论的。只能说明社会性线索起到了抑制作用，且社会性线索比感知觉线索具有更大的权重。

回应: 已修改，标注为黄色。

意见 6: 作者图中的误差线是标准差？一般来说，图中的误差线应该为标准误，以体现被试量的信息。

回应: 已将图中所有误差线改为标准误。

意见 7：一般来说，成人的注视时间平均为 300 ms 左右，儿童的平均注视时间可能比成人稍长一些那么，作者在此处剔除数据有何参考依据？

回应：评审老师，我不太理解您这里说的这句话：“一般来说，成人的注视时间平均为300 ms 左右，儿童的平均注视时间可能比成人稍长一些”。您是指一个注视点的平均注视时间吗？如果是一个注视点的平均注视时间，可能与本研究所指的注视时间不同。本研究的注视时间是指**整个视频的总注视时间**。理想状态下，整个视频的注视时间应该是整个视频呈现的时间。但由于存在被试注意水平的差异、眨眼、头动等偶然因素所引起的眼动数据的损失，因此每个视频的注视时间总是少于视频的呈现时间。那么，存在较大数据损失的试次将不具备很好的代表性，因此将作为奇异值予以删除。至于为何删除整个视频的注视时间少于1000ms的试次，是参考前人相关的研究（Fletcher-Watson, S., Leekam, S. R., Benson, V., Frank, M. C., & Findlay, J. M. (2008). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia*, 47(1), 248-257）。在他们的研究中，每个试次的呈现时间是3000ms，删除少于500ms的试次。而本研究由于每个被试做出选择的时间不同，因此每个试次呈现的时间分布在5000ms和8000ms内，平均呈现时间是6000ms左右，因此本研究将删除整个视频的注视时间少于1000ms的试次。

意见 8：一般来说，根据需要做交互作用在一个因素上的简单效应检验即可。下面也存在类似的问题。

回应：已修改，**标注为黄色**。

意见 9：表 2 包含了图 4 的信息，建议采用一种形式（图或者表）展示实验结果。下面也出现了类似的情况

回应：已删除了图。

意见 10：序号有问题？脸部注视时间的分析结果作者应该给出绝对量，而不只仅仅给出相对量（百分比）。

回应：序号已更正。**标注为黄色**。脸部局部注视时间分析的目的在于，考察AD儿童与TD儿童搜寻社会性信息时在空间分布上是否存在差异，并不关心他们在各个兴趣区注视时间的差

异。因为AD儿童由于注意力水平低于TD儿童，导致AD儿童整个视频的注视时间就显著少于TD儿童，这就进一步造成AD儿童在各个兴趣区的注视时间均显著少于TD儿童，而这种差异并不是本研究所关心的而是应该予以控制的额外变量，本研究所关心的是他们获取社会性信息的方式存在哪些差异，更多地是从哪个区域获取社会性信息，因此我们采用了各兴趣区注视时间比例进行分析。王乾东等人（2013）的研究中亦采用的是各兴趣区的相对量。由于不同年龄组被试看整脸的持续时间不同，并且为了控制眨眼或者眼睛瞬间脱移屏幕等导致被试总注视时间的不一致，因此他们也同样计算出注视眼睛、鼻子、嘴巴这些兴趣区的持续时间各自占注视整脸时间的比例。

意见11：换一种表述方式

回应：已修改。标注为黄色。

意见12：TD儿童好像也有类似的结果。

回应：已修改。标注为黄色。

意见13：因为被试取样的原因，审稿人认为，本研究的结果难以与前人的研究结果进行比较，即使进行了比较，也无法得出结果的不同就是年龄差异所导致的。

意见14：这段让审稿人感觉非常奇怪，从这里可以看出似乎作者要针对前人研究结果的争论提出本研究的问题，由于不同的前人研究采用不同生理年龄的ASD儿童进行研究，如果作者想要解决这一争议性的问题，应该操纵ASD儿童的年龄这一关键变量来进行研究。

而作者采用的ASD儿童的年龄与前人的研究存在差异，而且差异比较大，此外，本研究中ASD儿童与对照组的TD儿童在生理年龄上差别也比较大（前面已指出），这样的研究难以说清楚作者的研究结果与前人的研究结果差异的原因到底是什么？由于作者与前人的研究在多个方面存在差异，使得对实验结果的差异难以解释。

回应：关于本研究中ASD儿童与对照组TD儿童在生理年龄上的差异在批注3中已进行解释。此处不再赘述。

本研究与Gliga,等人（2012）、Parish-Morris等人（2007）和Hennon等人（2003）在实验设计上并没有差异。只是在已有研究的基础之上，增加了一个基线条件，用以回答已有研究中发现的一致条件下ASD儿童习得词语是否可能仅仅依据感知觉线索而并未觉察和利用社会性线索。同时，为了不仅仅给出最终的行为结果还要考察实时信息加工过程，我们还

运用眼动追踪技术考察 ASD 儿童是如何加工各种相关信息的做出最后的选择。我们将本研究与已有研究进行比较，主要是对矛盾条件的实验结果进行比较。而本研究所设置的矛盾命名条件与已有研究完全一致。因此，本研究在矛盾条件下的实验结果与已有研究结果存在的差异可以归因为被试选择的差异。因为，本研究所选取的被试是言语智力年龄与 6 岁儿童相当的学龄期 ASD 儿童，他们比已有研究中的被试具有更高的发展水平，因此在同样的实验任务中有更好的表现。本研究之所以要选取具有更高发展水平的 ASD 儿童，是想证实 ASD 儿童利用社会性线索习得词语的能力是能够得到发展的，他们不会一直停留在依据感知觉线索的婴幼儿早期发展阶段。正如许多 ASD 儿童共同注意的研究发现的那样，虽然大量研究表明共同注意缺陷是 1 岁之前诊断自闭症的主要证据(Jones & Carr, 2004), 但 ASD 儿童的共同注意只是发展上的延迟，并不会永久受损。他们达到一定的心理年龄后，就能对他人的眼睛注视方向或头部转向做出跟随反应(Leekam, Hunnisett, & Moore, 1998)。还有 ASD 儿童语言能力发展的研究也发现，虽然所有的 ASD 儿童在生命早期都存在语言发展障碍，但他们中的绝大部分在童年中后期都能够表现出一定的语言能力，尤其是在词汇发展方面(Hus, Pickles, Cook, Risi, & Lord, 2007; Sigman et al., 1999)。因此，我们选取的被试为具有一定语言能力的童年中后期的 ASD 儿童考察此类儿童是否掌握了较高水平的词语习得策略，是否并未停留在词语习得技能的早期发展阶段。为了说明本研究与已有研究的关系，作者对前言部分进行了重新梳理。修改部分标注为黄色。

评审老师还提到“如果作者想要解决这一争议性的问题，应该操纵 ASD 儿童的年龄这一关键变量来进行研究。”首先，作者并没有认为已有研究的结果存在争议，因为不同研究选取的被试不同，因此可能处于不同的发展阶段。依据词语习得的联合生成理论，可以很好的解释这种差异，因为处于不同发展阶段的儿童运用不同的词语习得策略，赋予不同的线索更高的权重。那本研究也是依据此理论，考虑选择具有更高发展水平的 ASD 学龄期儿童为被试的。其次，在 ASD 儿童的研究中想要严格操纵年龄变量来进行研究是很难实现的。因为，首先，ASD 儿童的数量非常有限；其次，ASD 儿童群体内部又存在极大的异质性，比如同样是生理年龄为 10 岁的两个 ASD 儿童，可能一个根本没有语言，可能另一个可能只是伴随一定的语言障碍，因此仅仅操纵 ASD 儿童的生理年龄是没有意义的。但是如果想要即控制其生理年龄又控制其智力发展水平，这是非常难实现的。就笔者目前所接触到的本研究领域内研究，大多数研究都是以智力发展水平与普通儿童进行匹配，而不将生理年龄作为变量进行操纵控制，作者目前还没有见到过一篇将年龄作为关键变量进行研究的文献。而本研究之所以要选取童年中后期的 ASD 儿童是考虑到，已有研究中只考察了具有较低言语智力

发展水平 ASD 儿童,而具有较高发展水平的 ASD 儿童是如何权衡感知觉线索和社会性线索,还不得而知。然而,只有处于童年中后期的 ASD 儿童才具有较高的发展水平,因此本研究才选取具有一定语言能力的处于童年中后期的 ASD 儿童为被试。

## 第二轮

审稿人 2 意见:

意见 1: 对取样的年龄差异问题解释比较充分。作者应该在正文中也做简要介绍。

回应: 已修改,标注为红色。

意见 1: 不是认为引用这篇文章不妥,个人觉得作者可以换个说法,而不是直接比较这两个研究,比如作者可以说:来自幼儿面孔识别领域的研究也发现。。。。。。从侧面支持了。。。。。(文献)

回应: 已修改,标注为红色。

意见 1: 统计符号应斜体,请修改文中部分。

回应: 已修改,标注为红色。

意见 1: 请作者在正文中也加以说明注视时间。在数据分析部分引用该文献,文献列表中添加该文献。

回应: 已修改,标注为红色。