

《心理学报》审稿意见与作者回应

题目：应聘者在人格测验中作假的反应过程：基于工作赞许性的眼动证据

作者：徐建平 陈基越 张伟 李文雅 盛毓

第一轮

感谢四位评审专家对本稿提出的宝贵审稿意见，也感谢评审专家们对本文的肯定。经过论文作者共同对审稿意见地仔细阅读，我们针对评审专家们的意见，查阅文献认真修改，逐条做出了解释以及修改说明（正文中红色字体内容为本次修改或增补内容）。具体如下，如有不妥当之处，敬请批评指正。

外审专家 1

该研究采用眼动方法考察了应聘者在人格测验中作假的反应过程，在前人研究的基础上增加了测题社会称许性的分类，研究具有理论意义和一定的实践价值。以下问题需要作者修改或思考：

1. 摘要中的最后一句“根据这一发现，提出基于工作称许性解释人格测验作假反应过程的混合加工模型。”因缺乏相应的内容，应去掉。

解释及修改说明 1：

根据评审专家的指正，已将摘要中的最后这句话删除。

2. 引言中的总体假设和具体假设需要调整。如总体假设“相比于诚实作答，工作称许性高的题目作假的反应过程更简单，工作称许性低的题目作假的反应过程更复杂……”的表述，对应于诚实作答的是作假作答，因而该句的第一个逗号后应有“作假作答”的字样，不仅如此，“反应过程简单”或“反应过程复杂”不易于理解，显得过于宽泛。此外，具体假设内容过多，且部分内容缺乏依据。

解释及修改说明 2：

根据评审专家建议，在总体假设中，将“作假作答”提到第一个逗号后面，与“诚实作答”形成对应。修改内容见第 3 页红色文字。

关于“反应过程简单”或“反应过程复杂”不易理解的问题，首先，因为本研究提出的假设是对前人关于人格测验作假三个理论的补充与完善，在前三个理论以及相关研究中，对比诚实作答与作假的反应过程的词语为“easy”与“difficult”，并通过项目反应潜伏期、眼动指标等对“easy”与“difficult”进行进一步阐述，因此作者们认为同样需要对本研究的假设进行一个总体的概述，沿用了前人的词汇，同时认为“difficult”一词如果直译为“困难”，不易理解，若翻译为“复杂”则与应聘者在作答时的认知思维这一过程更为贴近。因此本文作者们认为需要保留总体假设中对于诚实作答与作假作答的总体概述形容，并在后面的具体假设中通过反应潜伏期、眼动指标等对总体假设中的“简单”和“复杂”进行具体阐述。

关于“具体假设过多，且部分内容缺乏依据”的问题。本研究通过项目反应潜伏期、项目注视点个数，项目不同兴趣区注视点个数、眼动轨迹四个主要指标来检验应聘者在不同作答情境下的反应过程，因此对应了四个假设，这样不仅对总体假设中的“简单”与“复杂”进行了充分阐述，同时与结果部分的内容也形成了清晰的对应。关于部分假设内容缺乏依据的问题，作者在假设前的前言部分，先对三个理论模型进行了补充修改，之后又对三个理论模型有何不足，以及如何提出本研究的假设进行了补充，具体修改内容见第 2、3 页红色文字。

3. 该研究进行了智力测验，若该测验确实有必要，请阐述进行该测验的目的和意义，并呈现测验结果。否则建议删除。

解释及修改说明 3:

本研究为被试内研究设计，每名被试均需完成两次人格测验。在两次测验作答间隙加入了一个智力测验，其目的在于减少第一次作答测验对第二次作答测验产生的影响。该智力测验是一个干扰任务，有必要存在，但其智力测验分数对于本研究没有任何影响，因此没有呈现在文中。根据评审专家的建议，已修改对智力测验作用的描述，使其更加清晰明确，见第 4 页红色文字。

4. 该研究的实验设计呈现在程序中，且表述有误。建议更正实验设计的表述，并独立成段。

解释及修改说明 4:

根据评审专家的建议，作者阅读了《心理学报》近期发表的实验研究类论文，对实验设计的描述进行了修改，并将实验设计部分独立成段。见第 4 页“2.2 实验设计”红色文字。

5. 文中使用了注视点个数这一眼动指标，并在讨论中结合反应潜伏期进行了阐述，但没有使用注视时间的任何指标，请说明原因，并说明反应潜伏期、注视点个数、注视时间之间的关系。

解释及修改说明 5:

首先回答专家第一个问题，为何没有使用注视时间。本研究中，反应潜伏期记录的时间是指从一道题目出现至通过作答该题目点击鼠标这段时间，而眼动的注视时间是指在这道题目的反应潜伏期内，注视点注视时长的总和。本研究在初期结果分析的时候使用了注视时长这一指标，但经过分析发现，注视时长指标与反应潜伏期指标存在高相关，所有分析结果均与反应潜伏期指标的分析结果一致，因此没有使用注视时间这一指标。国外有关的研究中也仅使用了反应潜伏期、注视点个数指标，没有使用注视时间这一指标 (van Hooft & Born, 2012)。

专家提出的第二个问题，要求解释反应潜伏期、注视点个数、注视时间之间的关系。反应潜伏期与注视时间之间存在高相关，注视点个数与反应潜伏期、注视时间的关系存在两种情况：正常情况下注视点个数越多，反应潜伏期、注视时间越长，相关高；另一种情况，当被试注视点集中在一点时（可能集中精力思考问题，也可能走神），此时注视点个数少，但反应潜伏期、注视时间长。

6. 建议调整表3的表头名称，以体现“注视点个数”的字样。

解释及修改说明 6:

根据评审专家对表3表头名称表述问题的建议。现将原表3的表头名称“两种情境下注视点位置结果检验”改为“两种情境下不同位置注视点个数检验”，体现了“注视点个数”字样。修改内容见第8页表3表头红色文字。

外审专家 2

文章针对人格测验中作假的反应过程研究进行了实验研究，获得了一些有意义的结果。这是文章作者的贡献所在。但纵观全文，尚存在以下重要问题，需要进一步认真修改。建议编辑部根据修改情况决定是否录用或者退稿。

1. 作者对社会期许性和工作期许性的区分理论和经验的分析不够充分。特别是，人格作为一个基础性的心理学变量，不知作者如何对其做社会性与工作性的区分？

解释及修改说明 7:

社会赞许性和工作赞许性是有区别的。根据另外专家意见，本文将“称许性”、“期许性”全部修改为“赞许性”，所以后文回应专家意见和论文正文中均采用了这一名称。与基于社会称许的普遍性的作假行为不同，部分研究者主张人事选拔中的作假是一个特异的情境性行为。Sackett (2011)将受测者分数的变异划分为诚实分数、情境性偏差、一致的错误自我感知、情境性自我欺瞒、跨情境的印象管理、情境性的有意作假六个部分，并指出“情境性有意作假”反映了作假的真实内涵。在人事选拔情境中，应聘者会按照组织工作情境中的工作期望的标准来回答人格测验，倾向于向招聘方传递一种“优秀员工”形象，表现出来的是一种工作赞许性反应，而非单纯的社会赞许性(Bradley & Hauenstein, 2006; Mueller-Hanson, Heggestad, & Thornton, 2006)，比如一个“好公民”或“道德高尚的人”的形象。根据评审专家的建议，在引言中对有关工作赞许性和社会赞许性的区别和差异作了进一步描述和补充，见第1页红色文字。

2. 对有关三种作假反应模型的经验研究综述分析过于简单，因而在提出假设时过于仓促，知其然，不知其所以然。换言之，尽管作者给出似乎是比较清晰的假设；然而对于各个假设的根据作者并没有给出令人信服的理论解释。

解释及修改说明 8:

根据评审专家的建议，作者对综述中三种作假反应模型进行了补充，扩充了对于理论模型描述，同时增加了相关理论模型的实证研究结果，同时针对于本研究的假设，尤其是眼动部分，也增加了前人的研究结果的描述。增补内容见第2、3页红色文字。

3. 文题不太恰当。研究就是一些在校学生，如何跟“应聘者”联系在一起？

解释及修改说明 9:

感谢评审专家对文章题目的质疑。目前人事选拔的研究主要分为两大类，真实情境研究和模拟情境研究。本研究的被试虽然为在校大学生或研究生，但是采用的是模拟情境研究，

即模拟工作应聘的情境，被试模拟应聘者，研究的问题是应聘者在选拔情境下作答人格测验作假时的问题，因此在论文标题中包含了“应聘者”。为了明确该研究采用的是模拟情境，在中英文摘要中补充了相应自居，见第 1 页和第 14 页红色文字。

4. 作者在实验程序与情境的设置上是否存在不妥之处？文中的“作假”情境似乎不太恰当。

解释及修改说明 10：

本研究采用模拟情境研究，在国内外人格测验作假的模拟情境研究中，“作假”情境均是通过指导语进行的操作，已有了一套标准的流程。本研究严格按照人格测验作假模拟研究的实验设计进行，情境设置也符合标准。另附发表在 *Journal of Applied Psychology* 上的一篇关于人格测验作假模拟研究的文章信息供参考。

van Hooft, E. A., & Born, M. P. (2012). Intentional response distortion on personality tests: using eye-tracking to understand response processes when faking. *Journal of Applied Psychology*, 97(2), 301–316.

5. 文中尚存在一些明显的文字性错误。请作者认真校对。

解释及修改说明 11：

感谢评审专家对文中细节的纠正。几位作者已经轮流对文章中的字词句等文字错误进行了校对。

6. 英文摘要中的很多书写不符合英文习惯。请认真修改。比如，research 研究时，一般不用复数形式。

解释及修改说明 12：

感谢评审专家对英文摘要的纠正。作者已请英文专业老师一起对英文摘要进行修改，修改内容见第 14、15 页红色文字。

外审专家 3

该研究的主要创新性工作，是在考察应聘者在接受人格测验过程中作假反应模式时，基于工作称许性的概念，将测验题目分为高、低和无工作称许性三类，以解决现有作假解释模型(自我图式模型、语义加工模型和图式采择模型)间的冲突。该文的研究构想本身具有一定的新颖性，研究过程严谨规范。以下问题需要进一步澄清或商榷：

1. 作者认为，以往研究所以难以解决上述三种解释模型间的矛盾，是因为：(1) 这些模型均基于对社会称许性的研究，而应聘测验过程的工作称许性不同于一般的社会称许性；(2) 该类研究将测验题目的称许性分为高、低两种，忽视了被试对那些不具有很强的称许性项目的作假反应模式的考察(见“引言”第五段)。于是作者提出“是否存在另外一种基于工作称许性的反应过程模型呢？”的问题。审稿人对于第一点的论述没有异议，而对于第二点，审稿人认为：

第一，作者没有清楚地论证如下问题，即为何考察基于工作赞许性的作假反应模式，就必须同时考虑被试对无明显赞许性项目的作假反应模式？是否以往基于一般社会赞许性的作假反应研究也存在同样的局限？

解释及修改说明 13:

感谢评审专家提出的为何需要考虑无明显赞许性项目这一问题。首先在各种人格测验中，不论通过主观判断还是通过李克特量表评定，无明显赞许性的题目确实存在。而在前人的三种人格测验作假反应模型中，前两种没有对题目赞许性进行区分，第三种图式采择模型则对题目进行高赞许性和低赞许性之分，前人的理论模型确实忽略了无赞许性的题目。且在一些实证研究中，仅根据正反向计分来判断题目的赞许性属于高（正向计分）或低（反向计分），这种对于题目赞许性的评定方法与依据明显具有局限。因此，本研究将前人理论模型与实证研究中没有考虑到的无赞许性的题目放进研究中，试图得到更精确的研究结果，期望能在一定程度上解释前人三种理论模型互相矛盾的原因，并试图对人格测验作假反应模型从新角度做出解释。

第二，根据项目的赞许性程度，作者将项目分为高、低和无三类（见“引言”第六段）至少在中文语境中容易造成理解上的困难。根据“方法”部分的操作性定义，高赞许性（“2”）表示测验项目所表达的语义在工作中受赞许，低赞许性（“-2”）表示测验项目所表达的语义在工作中不受赞许，无赞许性（“0”）则表示测验项目所表达的语义在工作中受赞许或不受赞许无关。但是，在中文语境中，低赞许性可能不是“不受赞许”，而是被理解为“受赞许的程度较低”。此外，*desirability* 在国内通常被译为“赞许性”，赞许性未必是一个更佳的译法。据此，建议修改为受赞许项目、不赞许项目和无赞许性项目。

解释及修改说明 14:

感谢评审专家对于题目赞许性程度命名的修改建议。作者回顾了以往文献，发现对于赞许性高或低的题目，国外研究者通过“*desirable*”和“*undesirable*”来描述，确实如专家所说，不论从翻译角度或者中文语境理解角度，将其分为受赞许项目、不受赞许项目和无赞许性项目更加恰当。因此感谢专家提出的修改建议，现将全文关于题目分类的命名均改为受工作赞许项目、不受工作赞许项目和无工作赞许性项目，并将 *desirability* 按照国内通常翻译的名称，将“赞许性”、“期许性”修改为“赞许性”。修改内容见全文红色文字。

2. 由于人格测验中的无赞许性项目被定义为其内容无所谓被赞许或不被赞许，那么被试便不需要对该类项目做作假反应，正如该文所假设的，在主要测量指标上，在无工作赞许性的题目上，被试作假作答反应潜伏期与诚实作答反应潜伏期无显著差异。然而，在主要的行为和眼动指标上，对于无工作赞许性题目，诚实情境下的反应与作假情境下的反应却出现了显著或边缘显著的差异，因而与理论预测是不一致的。作者“讨论”部分没有对此进行深入讨论，却据此认为，“这一结果也与研究者关于作假自我图式模型的部分研究结果相似，在 DePaulo 等人(2003)的元分析中，发现作假与增加的反应潜伏期有关”（见“讨论”第2段），且“根据这一发现，提出基于工作赞许性解释人格测验作假反应过程的混合加工模型”（见“摘要”），不明白这个与假设不一致的结果如何能成为混合加工模型的证据。

解释及修改说明 15:

本研究假设在无工作赞许性的项目上，被试不会作假，因此在作假和诚实两种作答情境下，被试在题目反应潜伏期和眼动指标上无显著差异。研究结果显示，不论是反应潜伏期或眼动指标，在无工作赞许性的题目上，作假作答与诚实作答均出现了显著或边缘显著的差异，与研究假设不符，但结果与自我图式模型的结果是一致的。而在受工作赞许项目和不受工作赞许项目上，不论是反应潜伏期或眼动指标，作假作答与诚实作答的差异趋势是一致的，与假设部分相符，该结果与语义加工模型的结果是一致的。由于文中这部分的讨论，如评审专家所说，不太深入，因此对该部分讨论进行了增补，并进一步说明通过该结果是如何提出混合加工模型的。具体增补内容见第 10、11 页红色文字。

3. 由于在高工作称许性和低工作称许性的题目上，被试的作假反应模式与语义加工模型的解释一致，而“对于无工作称许性的题目，因为题目本身与工作称许性关联度不高，因此根据语义无法直接判断题目是好是坏，这时应聘者所采用的作答策略受到了阻碍，相对于诚实作答，作假的过程更加复杂”（“讨论”第五段），据此作为提出人格测验作假反应过程的混合加工模型。然而，作者没有对混合加工模型做出进一步的解释。例如，没有交代“混合”是谁与谁的混合，也没有指出对于无赞许性项目的加工是什么样的加工，只是说这个加工过程更加复杂而已。只所以如此，审稿人认为与“引言”部分对引入无赞许性项目的理论意义缺乏明确论证有关。

解释及修改说明 16:

感谢评审专家对本研究提出的混合加工模型这一作假解释模型构想的质疑。首先，在引文部分，研究引入无赞许性题目这一分类，这是基于在现实的人格测验中确实存在此类题目，且以前解释人格测验作假反应的三种模型没有考虑到这类题目，因此本研究加入了这一分类，并对其作假模式进行探讨，假设中认为因为无赞许性题目与工作情境无关，应聘者不会作假，因此在两种作答情境下各指标无明显差异。但是在实际研究结果中，发现在两种作答情境下，此类题目的各类指标都存在边缘或显著差异，且与自我图式模型的指标结果相一致。如评审专家所言，讨论中作者对混合加工模型缺乏详细的解释，也未指明“混合”是谁与谁的混合，也没有具体解释对无赞许性题目的加工是什么样的加工。在修改稿的讨论中对这些内容进行了补充和解释，具体修改见第 10、11 页红色文字。

注：本意见未涉及眼动研究的部分，此部分需要眼动研究专家把关。

综上，审稿人建议，该文大修后再审。

外审专家 4

近年来国外已有研究者将眼动技术用于应聘者在人格测验中的作假反应研究，本文以工作赞许性为切入点，也对此做了研究。研究选点明确，研究过程清楚，研究控制得当，积累了有价值的研究资料。值得商榷的一个核心问题是，工作赞许性本质上也是社会赞许性，虽然在应聘情境中有其特殊性——问题就是这种特殊性对于推进本领域的研究价值该

如何定义？如本文引言所言，目前研究者关于作假反应过程的解释存在三种不同的观点——自我图式模型，语义加工模型及图式采择模型，那么导致这种争议的关键是否是因为不同的情境？本文并没有在这点上立论，所得结果提出的所谓“混合加工模型”也就显得意义含混且价值不明。建议文章紧紧围绕工作赞许性这个选点做文章（相应的建议标题改为：“应聘者在人格测验中的作假反应：基于工作赞许性的眼动证据”），强化其在应聘情境中的独特性及其实际的和潜在的应用价值，至于理论推进还是谨言慎行。

解释及修改说明 17：

感谢评审专家对于文章主题的建议。作者同意评审专家的观点：工作赞许性本质上属于社会赞许性。但对第二位外审专家问题 1 的回答和解释中，作者也指出工作赞许性和社会赞许性也存在着区别。具体的区别描述见回答第二位专家第一个问题的内容。工作赞许性是指应聘者在组织工作情境或招聘选拔情境下，有意或无意地按照组织对优秀或合格员工的要求作为自己作答题目的一个判断标准，来回答测试题目，是一种基于工作要求而产生的作假行为。为了描述两者之间的差异，在此举一个例子。譬如“我的容貌很好”这一问题，大家都喜欢容貌姣好的人，这是人之常情。受社会赞许性的影响，对这一问题每个人也许都会做出赞许性回答，作答分数普遍高于实际状况。但在工作应聘情境中，如果应聘服务性行业，如酒店、餐饮、旅游、空乘服务这类职位时，因对容貌要求较高，应聘者比在一般情境下对此问题的回答更有可能产生作假行为，打分会更高；但对于申请生产一线的技术工、厨师等职位时，容貌并非这些职位必须的要求，应聘者可能不会作假或过多思考，可能会按照一般情境下受社会赞许性的影响作答，也有可能按照自己容貌真实情况如实作答，因此该类应聘者对于这种题目打高分的可能性和高分的程度就不会高于申请服务行业的应聘者。当然，这只是一个现象性描述，需要研究去验证。正文中对于工作赞许性的补充修改内容见第 1 页红色文字。

对于以往的三种理论模型，作者从两个方面进行了探讨，一是提出基于工作赞许性进行研究，二是相对于图式采择模型中把题目分为受赞许和不受赞许两类这一作法，又区分出无赞许性题目这一类。实际上本文的立论与评审专家提出的“建议文章围绕工作赞许性这个点做文章”是吻合的。对于评审专家提出的“混合加工模型”意义含混的问题，作者在讨论部分也进行了修补。对于评审专家关于理论推进需要谨言慎行的意见，在最后的讨论中指出本文根据研究结果，综合三种理论模型，提出的这一种新的模型构想，只是一个探索和尝试，尚待以后进一步研究和验证。修改内容见第 11 页红色文字。

最后，为了紧紧围绕工作赞许性，作者采纳评审专家的建议，以突出研究在应聘情境中的独特性及其实际和潜在的应用价值，并将标题改为“应聘者在人格测验中作假的反应过程：基于工作赞许性的眼动证据”，修改内容见第 1 页和第 14 页标题红色文字。

第二轮

感谢三位评审专家对本稿提出的宝贵审稿意见，也感谢评审专家们对本文的肯定。经过论文作者共同对审稿意见地仔细研读，我们针对评审专家们的意见，查阅文献认真修改，做出了解释以及修改说明（正文中绿色文字内容为本次修改或增补内容）。具体如下，如有不

妥当之处，敬请批评指正。

外审专家 1

作者对文章做出了认真、合理的修改，建议发表。

外审专家 2

文章针对人格测验中作假的反应过程研究进行了实验研究，获得了一些有意义的结果。这是文章作者的贡献所在。但纵观全文，尽管作者针对审稿专家的意见进行了认真和必要的修改，有了明显的改进。但本人认为尚存在一些小问题，需要作者进一步在理论和方法上加以澄清。

1. 文中关于眼动指标方面，不知作者是否考虑到注视点位置和眼动轨迹中可能出现优势眼偏差问题？请作者加以澄清。

解释及修改说明 1:

感谢专家的宝贵意见。本研究在实验过程中，考虑了优势眼、色盲、色弱、高度近视等可能影响记录结果的因素，然后一一筛查，做到对入组被试的控制。对于色盲、斜视和色弱，主要是通过询问被试，如有则不能参加本实验；近视者如矫正视力正常，先进行练习实验，根据练习实验记录的眼动模式，决定是否让其参加正式实验。如果练习实验记录正确，眼动模式正常，则继续参与正式实验。对于被试的优势眼，研究者制定实验方案时已经考虑到，了解了优势眼简易测试方法，查阅了如何在眼动实验中减小优势眼的偏差。

在本实验中，主要是从以下几个角度尽可能来减少优势眼的偏差：

(1) 询问被试。实验前询问被试是否有优势眼。如果被试不理解优势眼，实验人员就给予详细解释，并告诉被试的建议检测方法。如果被试的左右眼差异很大，该被试就不能参加下一步的实验；

(2) 分析校准结果。在询问被试基础上，进行练习实验，通过实验前的校准结果来监测被试左右眼的差异。如果校准结果中左右眼的差异很大，即一只眼的结果与另一只眼的结果存在系统性偏移，那么该被试就不能进入后面的正式实验。

(3) 通过仪器修正。对于所有进入正式实验的被试，本实验中使用的 Tobii 眼动仪记录眼动时采用的是“双眼平均记录法”，眼动仪同时记录两只眼睛的运动情况，然后对两只眼的值进行平均，最后得出被试的注视位置和注视时间等眼动指标。论文中报告的眼动指标结果均为对被试双眼记录平均后的结果。Tobii 眼动仪的算法是通过左、右眼的 Gaze 坐标进行平均计算出 mapped gaze 坐标值，这一过程由眼动分析软件自动完成。平均的过程是基于前期校正偏差的数据来进行，自动去除了双眼偏差，尽可能达到精确记录。经过处理的眼动数据最终是叠加在刺激材料上并在刺激材料坐标系中以像素为单位计算的，无论哪只眼睛是优势眼（或特殊情况如斜视），这就是被试真实看到的位置，即有效的注视点。详细可参考 tobii T120 操作手册、说明手册（www.tobii.com）以及 tobii 注视点算法手册中的 3.1.2 部分（The Tobii I-VT Fixation Filter Algorithm description）。

(4) 在数据处理阶段，研究者者在划分兴趣区时充分考虑到了知觉广度和不可避免的记录偏差，使更多有心理意义的数据能包含到兴趣区中。

具体修改内容见正文第 3、4 页“2.1 被试”和“2.3 仪器”中绿色文字。

参考文献: Olsen, A.. (2012). The Tobii I-VT Fixation Filter Algorithm description.

2. 文章篇幅是否过长? 在修改稿中, 作者努力照顾审稿专家的意见, 加入了不少新的内容。但很多地方过于冗余拖沓。建议作者对在修改稿中加入的修改内容再进行必要的浓缩精炼, 力求做到言简意赅。

解释及修改说明 2:

感谢评审专家对文章内容及篇幅提出的修改建议。作者们已根据专家的建议, 对全文文字进行了浓缩精炼, 删除、缩减了冗余字句, 修改见文章中的绿色文字。

外审专家 3

该文修改稿对一审中提出的问题做了较好的回应, 并做了相应的修改, 澄清了相关的问题, 总体质量有了明显提高。建议刊用。