

复杂度，准确度和流利度对于高考英语和托福写作分数占比影响分析

曾 焱

新东方国际教育培训事业部，重庆

收稿日期：2021年11月5日；录用日期：2021年11月23日；发布日期：2021年11月30日

摘 要

复杂度，准确度和流利度是二语习得领域用来衡量语言发展水平的三个重要语言特征，可以用来指导不同考试的备考。本研究通过分析高考英文写作和托福写作的评分标准和写作数据，探讨复杂度，准确度和流利度对于最终分数的影响情况。

关键词

复杂度，准确度，流利度，第二语言习得，高考，托福，语言测试

How Complexity, Accuracy, and Fluency Shape Scores of English Writing in National College Entrance Examination and TOEFL

Yan Zeng

Global Education and Training Division, New Oriental Education & Technology Group, Chongqing

Received: Nov. 5th, 2021; accepted: Nov. 23rd, 2021; published: Nov. 30th, 2021

Abstract

Complexity, accuracy and fluency (CAF) are three important predictors that are used to describe second language performances in the field of second language acquisition (SLA). Therefore, they could be utilized to help with the preparation of varying tests. By analyzing writing rubrics and actual writing data, this study discusses the influence complexity, accuracy and fluency exert on writing section from both National College Entrance Examination and TOEFL.

文章引用：曾焱. 复杂度，准确度和流利度对于高考英语和托福写作分数占比影响分析[J]. 国外英语考试教学与研究, 2021, 3(4): 158-168. DOI: 10.12677/oetpr.2021.34018

Keywords

Complexity, Accuracy, Fluency, SLA, National College Entrance Examination, TOEFL, Language Testing

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

为了描述第二语言能力，二语习得领域重点强调三个语言特征来衡量语言发展的质量：复杂度、准确度和流利度[1]。但是国内考生参加的英语考试对于这三个方面的偏向程度却差别非常大。所以，无论对于老师还是对于学生来说，需要了解该考试在这三个方面的不同要求，这样学生才能有的放矢，针对性的获得理想的成绩。本文将通过对比国内高考英语最后一题作文和托福考试中的独立写作单项要求，分析三个因素在二语习得的占比影响，帮助学生更好的备考高考英语或者准备出国英语考试。

2. 复杂度，准确度和流利度的定义

2.1. 复杂度

语言学家 Ahmadian 给语言的复杂度下的定义是这样的：“学习者想要冒险并且使用他们最前沿的语言知识的倾向”[2]。虽然复杂度涉及的方面非常多，但大家最常提及的，还是句法，词汇这些语言方面的复杂度。比如说，词/句长，并列和从属句式的复杂度，以及其他的复杂语法结构。

但根据研究，复杂度和语言能力的相关度并不是很高[3]。此外，复杂度还面临着主观性非常强的挑战，即在不同人的眼中，复杂度的高低是不同的。比如，同一个词汇在中国老师的眼中也许会得到很高复杂度的评价，而在母语人眼中却是难度一般的使用。因为每个级别涉及到的词汇量非常大，所以在主观评分时，很难当下准确的判断语言的复杂度。

2.2. 准确度

在写作中，准确度的考察是通过测量错误的数量来决定的。具体来讲，就是作文中使用的单词、短语、句子中有错误和没错误的比例来决定一篇文章的准确度高低。此外，还有评分方式的影响，国内的考试(例如高考)一般采用分析性评分原则，会针对语言单项错误进行扣分；而托福采取整体评分原则，估算大概的错误比例。

准确度的研究，以往语言学家也做了很多，都呈现出和语言水平直接相关的结论[4][5]。在国内考试和托福考试中，准确度也都和分数的关系非常紧密，分数低的作文比分数高的作文准确性是要更差的，这一特征尤其在低分段和中分段中更为常见。

2.3. 流利度

流利度是平时大家接触最少的一个语言特征。事实上也是国内英语考试和托福考试最大的一点区别。我们先来看它的定义。Wolfe-Quintero 把流利度定义为语言提取的自动反应速度[6]。它是用来衡量一个人实时的思考、组织和传递语言的能力的重要指标。

对于写作来说，流利度体现在同样的限时时间内，产出文章长度的不同。语言能力更好的同学能够比语言能力逊色的同学写出更长的文章来，所以字数就是一个非常直观的衡量第二语言学习者流利度的指标。

2.4. 复杂度，准确度和流利度之间的关系

总体来说，复杂度，准确度和流利度(以下简称“三度”)既是相辅相成的，又是相互制约的。首先，在写作的过程中不可能把这三个要素刻意分开，是有机结合在一起共同完成文章的；但另一方面，这三个要素又是相互影响的。比如，如果一位学生想要追求写作准确度，那么他就要牺牲速度，多用些时间来检查。相反，如果一位学生想要写得多，写得快，那么自然他的写作准确度就会受到影响。复杂度亦是如此。例如，如果一位学生想要追求写作复杂度，就会不可避免的牺牲一部分准确度。同样的，追求复杂度也意味着不可能流利度同时做到非常好。

3. 写作评分标准和复杂度，准确度，以及流利度的关系

3.1. 高考和托福写作评分标准

无论是国内的高考，还是托福考试，写作的评分维度是一样的，都是三个角度：内容，语言和组织。以下是 2021 年全国高考卷作文的评分标准，可以明确看出对于三个评分角度不同档次的要求：

各档次给分要求

内容部分

- 1) 内容充实，主题突出，详略得当。
- 2) 内容基本充实，尚能表达出作文要求。
- 3) 漏掉或未能写清楚主要内容，有些内容与主题无关。
- 4) 明显遗漏主要内容，严重离题。

语言部分

- 1) 具有很好的语言表达能力，语法结构正确或有些小错误，主要因为使用了较复杂结构或词汇所致。句子结构多样，词汇丰富。
- 2) 语法结构和词汇方面有错误，但不影响理解。句子结构有一定的变化，词汇使用得当。
- 3) 语法结构与词汇错误较多，影响了对内容的理解。句子结构单调，词汇较贫乏。
- 4) 语法结构与词汇的错误很多，影响了对内容的理解。词不达意。

组织结构部分

- 1) 自然地使用了语句间的连接成分，全文流畅，结构紧凑。
- 2) 能使用简单的语句间连接成分，全文内容连贯。
- 3) 尚能使用语句间连接成分，语言连贯性较差。
- 4) 缺乏语句间的连接成分，语言不连贯。

托福官方指南第 6 版的 P203-P204 也给出了托福考试的写作评分标准(见附录 1) [7]，也是三个角度 Development, Organization and Language Use，和分别和高考评分角度“内容部分”，“组织结构部分”和“语言部分”呼应。

3.2. 写作评分角度和复杂度，准确度，以及流利度的关系

这三个评分角度和三度是三三对应的关系，具体来讲：

第一，内容。内容的复杂度指的是内容的使用上选择一些不常见的、读者理解度较低的事例。比如

同样论证创新很重要，如果内容举例是爱迪生发明灯泡，复杂度就低一些；而写卡尔·冯·克劳塞维茨创作出了《战争论》，复杂度就高一些。当然，正如前文所说，复杂度仍然还是主观性非常高的一项指标。内容的准确度指的是内容是否是正确的，而和写作篇幅并不直接相关。与之对应的，就是内容的流利度，指的是展开的充分程度。同一个主旨下，展开 3~5 句话，还是 8~10 句话。所以，内容流利度最为相关的一个指标，就是文章的字数。

第二，组织。组织结构的复杂度指的是是否运用各种手段体现内容与内容之间的联系。比如在篇章层面，没有采取正序，而是采取倒序或者插叙的形式来写作，复杂度就更高一些。在段落层面，除了有表面的词句连贯性外，是否还有一条隐含的语意逻辑性。在句子层面，除了用连接词，是否还用到其他不常见的连接方式，比如用代词等。组织的准确度即逻辑关系是否正确，比如是否举例跑题，内容重复，或者前后文是转折关系，作者却处理成因果关系等等，这样可能导致读者不能正确理解文章逻辑。最后，组织的流利度，指的是写作者处理各种逻辑关系的反应速度，可以通过限时写文章中的逻辑关系使用数量来衡量。

第三，语言。这是上文讨论最多，也是一般理解三度最先触及的部分。语言的复杂度指的是句法，词汇等复杂度。准确度指的是语言错误占全篇文章的比例。流利度指的是大脑中反应某个词，运用某个句型和语法的反应速度，也可以体现在字数上。

4. 高考写作要求分析

4.1. 高考写作评分标准对于三度的要求

针对高考满分作文评分标准的描述，将相关证据填到表格中，对应结果如下(表 1):

Table 1. The correlation between rubrics of English writing in National College Entrance Examination and complexity, accuracy and fluency
表 1. 高考英语写作评分标准和三度的对应

	复杂度	准确度	流利度
内容	低(给定话题，限定内容作文)	高(内容充实，主题突出，详略得当)	低(总字数为 120~150 词，但时间无单独限制)
组织	中(自然地使用了语句间的连接成分)	高(全文流畅，结构紧凑)	低(时间无单独限制，无法测量)
语言	高(句子结构多样，词汇丰富)	高(具有很好的语言表达能力，语法结构正确或有些小错误，主要因为使用了较复杂结构或词汇所致)	低(时间无单独限制，无法测量)

从评分标准的解读来看，如果一名学生想要在高考写作中获得满分的话，重要性排名最重要的指标是“准确度”。内容，组织和语言三个角度都在这个方面要求很高。排名的第二的是“复杂度”，除了内容是限定的，复杂度要求较低外，对于组织和语言的复杂度都有较高要求。尤其是语言，满分作文中的语言要求复杂度比准确度更为重要。正如满分标准所言，“语法结构正确或有些小错误，主要因为使用了较复杂结构或词汇所致”。

所以，从评分标准中可以推断出，三度的重要性排名为：准确度>复杂度>流利度。即，如果学生想得到满分的话，首先他要准确度特别高，其次复杂度，尤其是语言的复杂度也要很高。即使因为写的很复杂，牺牲一些准确度，仍然可能得到满分。相反，流利度是没有考察(没有单独限时)，并且严格限制了

的(不能超过 150 词)。那实际操作中, 准确度和复杂度能否成为评判学生作文分数的指标呢?

4.2. 研究过程

根据上海 2020 年高考二模中的英文写作题目(满分 25 分, 见附录 2), 一位参与高考阅卷评分人提供的 30 份成绩单, 这些成绩单的分数分布如表 2 所示。本文从他们的分数和三度的关系来进行了分析。

Table 2. Distribution of scores

表 2. 样本分数分布

分数	15	16	18	19	20	22	23	24	总计
数量	1	3	4	4	7	6	3	2	30

在这 30 篇作文当中, 我统计了他们的错误个数、字数、以及正确率, 如表 3 所示:

Table 3. Descriptive statistics for variables across score levels of English writing in National College Entrance Examination

表 3. 高考英语作文样本错误率统计

编号	语言错误个数	字数	错误率	分数
1	1	148	0.68%	24
2	2	162	1.23%	24
3	3	170	1.76%	23
4	3	155	1.94%	23
5	4	141	2.84%	23
6	5	148	3.38%	22
7	7	144	4.86%	22
8	7	152	4.61%	22
9	6	153	3.92%	22
10	8	150	5.33%	22
11	6	145	4.14%	22
12	10	144	6.94%	20
13	12	169	7.10%	20
14	13	166	7.83%	20
15	11	140	7.86%	20
16	11	155	7.10%	20
17	12	145	8.28%	20
18	14	130	10.77%	20
19	14	136	10.29%	19

Continued

20	16	125	12.80%	19
21	16	119	13.45%	19
22	17	128	13.28%	19
23	20	169	11.83%	18
24	21	148	14.19%	18
25	19	153	12.42%	18
26	20	134	14.93%	18
27	25	123	20.33%	16
28	28	155	18.06%	16
29	26	149	17.45%	16
30	30	120	25.00%	15

其中分数区间为 15 分~24 分, 错误数量区间为 1~30 个, 字数区间为 120~170 词。我们再来通过图表的形式看一下错误数量, 字数和最终分数的关系。

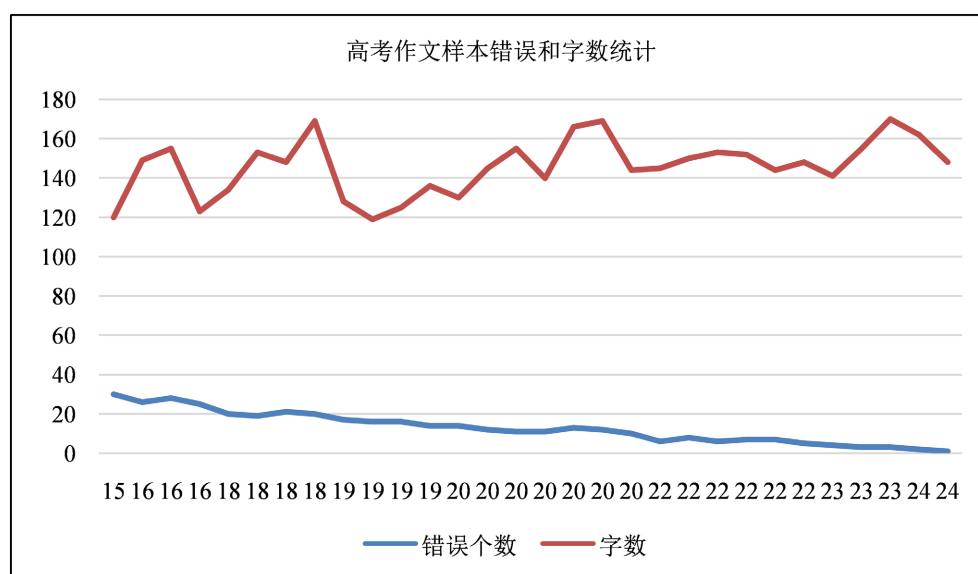


Figure 1. Accuracy and word count across score levels of English writing in National College Entrance Examination

图 1. 高考英语作文样本错误和字数统计

由图 1 可以看出, 学生样本的字数统计并没有和分数呈现很强的相关性, 这可能是因为他们都满足了考试要求的基本字数, 所以在这个方面并没有对分数造成任何影响。而语言错误个数的趋势非常明显, 尽管有小的波折, 但整体趋势都是错误越少, 得分越高。错误率的分布规律基本和错误个数相似。由此, 我们可以得出结论, 高考作文分数和准确度是非常密切相关的, 而和流利度(字数)没有什么相关性。

此外, 我也统计了这 30 篇文章的词汇复杂度和分数的关系, 如表 4 所示:

Table 4. Word complexity across score levels of English writing in National College Entrance Examination
表 4. 高考英语作文样本词汇复杂度统计

编号	% NGSL 1K	% NGSL 2K	% NAWL	分数
1	68%	23%	9%	24
2	65%	26%	9%	24
3	68%	23%	9%	23
4	72%	19%	9%	23
5	71%	21%	8%	23
6	72%	20%	8%	22
7	72%	20%	8%	22
8	73%	19%	8%	22
9	74%	18%	8%	22
10	74%	18%	8%	22
11	74%	18%	8%	22
12	75%	17%	8%	20
13	76%	16%	8%	20
14	74%	19%	7%	20
15	75%	18%	7%	20
16	75%	18%	7%	20
17	75%	18%	7%	20
18	75%	18%	7%	20
19	75%	18%	7%	19
20	77%	16%	7%	19
21	77%	16%	7%	19
22	80%	13%	7%	19
23	81%	14%	5%	18
24	79%	16%	5%	18
25	77%	18%	5%	18
26	77%	18%	5%	18
27	85%	10%	5%	16
28	87%	8%	5%	16
29	85%	10%	5%	16
30	92%	5%	3%	15

NGSL 是语言学家 Dr. Charles Browne, Dr. Brent Culligan, Dr. Joseph Phillips 的研究成果。他们通过分析剑桥英语语料库,总结了在普通英文文本中出现频率最高的前 2800 个单词。其中前 1K 更为高频。NAWL 是语言学教授 Averil Coxhead 通过研究语料库,统计得出在所有的学术类文章中出现频率最高的 960 个英语单词。他们要难度更高一些,词频也比 NGSL 里面的词更低一些。

其中 NGSL 1K 区间为 68%——92%，NGSL 2K 区间为 5%——23%，NAWL 区间为 3%——9%。我们再来通过图表的形式看一下词汇复杂度和最终分数的关系，更加的直观。

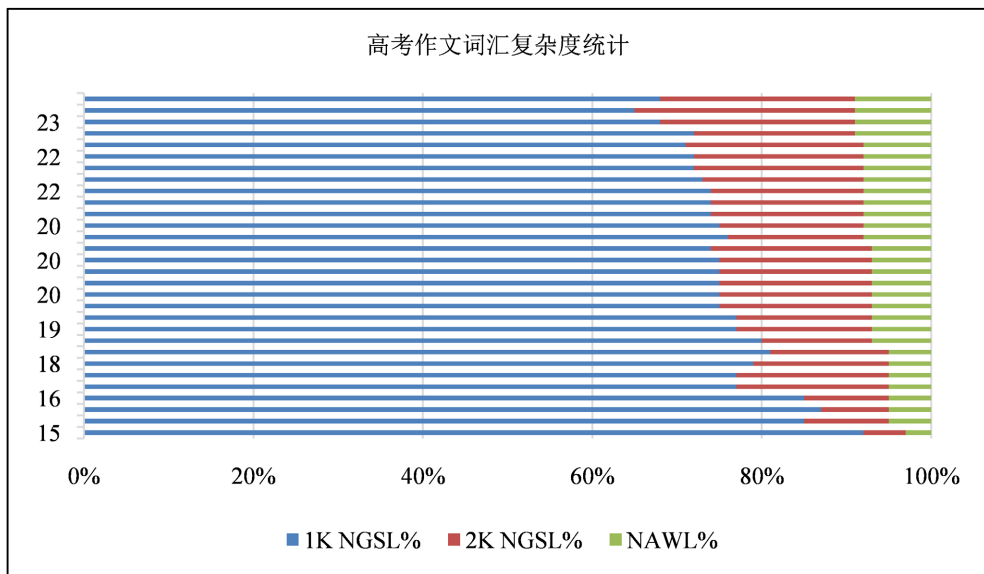


Figure 2. Word complexity across score levels of English writing in National College Entrance Examination

图 2. 高考英语作文词汇复杂度统计

由图 2 可以看出，大体而言，随着学生分数的上涨，NGSL 1K 的比例会下降，而 NGSL 2K 和 NAWL 的比例会上升。这表明随着复杂度的提高，确实分数会更高。这也和高中学校老师要求多积累“好词好句”，“多背范文”的教学方法和理念是一致的。

总之，通过这 30 份的数据分析，基本证实了评分标准的说明，即：评分是根据错误数量多少来扣分的，错误的数量越多，分数就越低。此外，高分同学的词汇复杂度要比低分同学更好。而字数，即流利度和最后的成绩相关性并不高。

5. 托福独立写作要求分析

针对托福独立写作满分作文评分标准的描述，将相关证据填到表格中，对应结果如下(表 5)：

Table 5. The correlation between rubrics of TOEFL writing and complexity, accuracy and fluency

表 5. 托福独立写作评分标准和三度的对应

	复杂度	准确度	流利度
内容	低(对于内容并无任何限制，也没有任何复杂度要求)	中(有效的回应了题目和任务)	高(展开充分的，运用了恰当的解释，举例和细节)
组织	低(没有对复杂度提出要求)	高(统一，递进，连贯)	高(30 分钟，至少 300 词)
语言	中(句式多样性)	中(可以有词汇或者语法上的小错误)	高(30 分钟，至少 300 词)

从评分标准的解读来看，如果一名学生想要在托福写作中获得满分的话，重要性排名最重要的指标是“流利度”。内容，组织和语言三个角度都在这个方面要求很高。排名第二的是“准确度”，基本上

要求可以有各种各样的小错误，但不影响理解就可以了。“复杂度”排名第三，并且只在语言的复杂度方面有要求。

所以，从评分标准中可以推断出，三度的重要性排名为：流利度 > 准确度 > 复杂度。即，如果学生想得到满分的话，他准确度不需要特别高，可以有各种小错误，只要不影响理解就可以。而流利度需要做到很好，体现在语言产出快，文章字数多，而且字数要多在发展的部分，而不是多在开头结尾，因为在内容展开方面是要求了流利度的。相反，复杂度是考察要求最低的一环，并且只要求了语言的复杂度，没有要求内容和逻辑的复杂度。这也就是教师平时在教学中常说的，内容不用“名人文学历史”，而用“日常生活例子”；逻辑关系只要用逻辑关系词和信号句能处理好就可以了，不用使用各种连接手段。那实际操作中，流利度和准确度能否成为评判学生作文分数的指标呢？

这方面 ETS 已经做过很多研究，这里引用一篇论文的结论。1) 流利度是写作水平最有力的信号；2) 语法和词汇准确度将会随着分数上升而减少；3) 复杂度有一个重要的，但是相对较小的影响[8]。

6. 结论

相比较而言，高考和托福考试的写作评分重点非常不一样。高考英语写作的评分重点是在准确度很高的基础上，追求复杂度才能获得高分，设计方式和第一语言测评方法接近。而托福考试更符合为第二语言设计的考试，建立在以测试考生交流能力为目的的基础上，所以流利度是最为重要的指标。准确度虽然也重要，但满分作文仍然允许写作者犯很多的小错误，只要它们不影响交流就可以了。而对于过高的复杂度，反而会影响交流，而不被托福考试所推崇。

知道这个结论，对于教授托福考试的教师来说，就知道了学生在接受托福培训前按照高考考试要求学习的习惯是怎样的。他们的优势在于准确度基础还不错，但劣势在于过度注重复杂度而不在意基本交流的达成。此外，流利度的缺乏也是几乎所有学生的标签。明白了这一点，教师在教学中就更能提高效率，有的放矢。

参考文献

- [1] Housen, A., Kuiken, F. and Vedder, I. (Eds.) (2012) Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA. <https://doi.org/10.1075/llt.32>
- [2] Ahmadian, M.J. (2011) The Effect of “Massed” Task Repetitions on Complexity, Accuracy, and Fluency: Does It Transfer to a New Task? *The Language Learning Journal*, **39**, 269-280. <https://doi.org/10.1080/09571736.2010.545239>
- [3] Neumann, H. (2014) Teacher Assessment of Grammatical Ability in Second Language Academic Writing: A Case Study. *Journal of Second Language Writing*, **24**, 83-107. <https://doi.org/10.1016/j.jslw.2014.04.002>
- [4] Polio, C. (1997) Measures of Linguistic Accuracy in Second Language Writing Research. *Language Learning*, **47**, 101-143. <https://doi.org/10.1111/0023-8333.31997003>
- [5] Polio, C. (2001) Research Methodology in Second Language Writing Researching: The Case of Text-Based Studies. In: Silva, T.J. and Matsuda, P.K., Eds., *On Second Language Writing*, Lawrence Erlbaum, Mahwah, 91-115.
- [6] Wolfe-Quintero, K., Inagaki, S. and Kim, H.Y. (1998) Second Language Development in Writing: Measures of Fluency Accuracy and Complexity. University of Hawaii at Manoa, Honolulu.
- [7] ETS (2021) The Official Guide to the TOEFL iBT Test, 203-204.
- [8] Plakans, L., Gebril, A. and Bilki, Z. (2019) Shaping a Score: Complexity, Accuracy, and Fluency in Writing Performances. 175.

附录 1：托福独立写作评分标准

5 分：

An essay at this level largely accomplishes all of the following:

- Effectively addresses the topic and the task
- Is well organized and well developed, using appropriate explanations, exemplifications, and/ or details
- Displays unity, progression, and coherence
- Displays consistent facility in the use of language, demonstrating syntactic variety, appropriate word choice, and idiomaticity, though it may have minor lexical or grammatical errors

4 分：

An essay at this level largely accomplishes all of the following:

- Effectively addresses the topic and task well, though some points may not be fully elaborated
- Is well generally organized and well developed, using appropriate and sufficient explanations, exemplification, and/or details
- Displays unity, progression, and coherence, though it may contain occasional redundancy, digression or unclear connections
- Displays facility in the use of language, demonstrating syntactic variety, appropriate word choice, and range of vocabulary, though it may have occasional noticeable minor errors in structure, word form, or use of idiomatic language that do not interfere with meaning.

3 分：

An essay at this level is marked by **one or more of the following**:

- Addresses the topic and task well using somewhat developed explanations, exemplification, and/or details
- Displays unity, progression, and coherence, though connections of ideas may be occasionally obscured
- May demonstrate inconsistent facility in sentence formation and word choice that may result in lack of clarity and occasionally obscure meaning.
- May display accurate but limited range of syntactic structures and vocabulary

2 分：

An essay at this level reveals **one or more of the following** weaknesses:

- Limited development in response to the topic and task
- Inadequate organization or connection of ideas
- Inappropriate or insufficient exemplifications, explanations, or details to support or illustrate generalizations in response to the task
- A noticeably inappropriate choice of words or word forms
- An accumulation of errors in sentence structure and/ or usage

1 分：

An essay at this level is seriously flawed by one or more of the following weaknesses:

- Serious disorganization or underdevelopment
- Little or no detail, or irrelevant specifics, or questionable responsiveness to the task
- Serious and frequent errors in sentence structure or usage

0 分：

An essay at this level merely copies words from the topic, rejects the topic or is otherwise not connected to

the topic, is written in a foreign language, consists of keystroke characters, or is blank.

附录 2：2020 年上海高考二模写作

Directions: Write an English composition in 120~150 words according to the instructions given below in Chinese.

假设你叫王林，你的美国好友 John 来信说他最近每天晚上玩手机到深夜，导致白天没有精神且食欲不振。他对此很担心，但又管不住自己，因此写信向你求助。请你用英语给 John 写一封回信，谈谈你的想法并给他一些建议。