財團法人語言測驗訓練中心
外語教學、測驗研究補助專案

結案報告

申請人	林銘輝
就讀學校及系所	淡江大學英文學系
研究計畫名稱	英語寫作語法複雜度、正確度、以及流暢度之抵換關係與寫作表現：全民英檢中高級寫作測驗之研究
計畫執行期間	2014年10月－2015年9月

摘要

「有限注意力容量配置」（Limited Attentional Capacity Model，簡稱 LACM）（Skehan & Foster, 1997, 2001）能影響大腦處理記憶資訊的進行，造成以英語為外語（English as a Foreign Language，簡稱 EFL）的寫作複雜度（Complexity）、正確度（Accuracy），以及流暢度（Fluency）（CAF）三者之間發生不同程度的抵換關係（Trade-off Relationship）。在眾多不同的因素中，時間可謂左右 EFL 作家對 LACM 掌控的重要變項之一，然而實際上專研「限時寫作」（timed writing）中 CAF 互動表現的文獻並不多，這樣的情況在台灣 EFL 寫作研究中尤其缺乏。此外，以限時寫作樣本為文本，探究不同寫作能力與 CAF 互動關係的文獻更是匱乏。為能豐富此一專門領域知識，本計畫檢視 150 份全民英檢中高級考生的寫作樣本，以量化統計與敘述統計進行分析研究。結果顯示，台灣 EFL 全民英檢考生不論其寫作能力高低，限時寫作的表現確實反映 LACM 理論，顯示在時間有限的條件下，複雜度與流暢度/正確度之間有抵換關係：文章之複雜度愈高，流暢度/正確度愈低，而流暢度愈高，正確度愈高，但以上的抵換關係與分數的相關性未達顯著水準。本報告並將對這樣的結果進行深入討論，並建議未來可行之研究方向。

關鍵詞：有限注意力容量配置、抵換關係、英語外語寫作

Abstract

The Limited Attentional Capacity Model (LACM) (Skehan & Foster, 1997, 2001) is believed to have an impact on the memory processing system, leading to a trade-off model between the levels of complexity, accuracy, and fluency (CAF) in writing output by EFL learners. Among the various factors, time may be one critical variable consuming EFL writers' LACM, but the current literature on the possible influence of LACM on the CAF relationship demonstrated in timed writing is rather limited; this is particularly the case in the context of Taiwan. What has not been properly investigated here includes the investigation of possible CAF interaction caused by Taiwanese EFL writers' proficiency levels. Studying this is also important given how the development of student writers' skills change over the course of learning, that is, being able to write with more complexity, more accurately, or more fluently when they are more skilful in writing than when they are not. To shed light on these aspects, the researcher of this study investigated 150 timed writing samples created by EFL test-takers who participated in the General English Proficiency Test High Intermediate. Both inferential and descriptive statistics were used to analyze the data. The research results indicate a possible trade-off relationship between complexity and fluency/accuracy, irrespective of the overall writing proficiency of the GEPT test-takers. To be specific, the higher the complexity of the writing samples, the lower their fluency and accuracy; the higher the fluency of the samples, the higher their accuracy. However, such trade-off relationship is not significantly correlated with the test-takers' writing scores. The paper concludes by offering implications and by identifying possible avenues for further studies.

Keywords: the Limited Attentional Capacity Model (LACM), EFL writing

INTRODUCTION

Over the past decades, writing complexity, accuracy, and fluency (also known as CAF) have received much attention in the field of second language acquisition (SLA), particularly in the field of English as a foreign or second language (EFL/ESL) (e.g., Ellis & Yuan, 2004; Lin & Chen, 2015; Lin, Chen, & Chen, 2015; Ong & Zhang, 2010; Polat & Kim, 2013; Robinson & Gilabert, 2007; Vardazaryan, 2012). While many investigators confirm the validity of CAF as effective indicators of language learners' performance[1] (cf. Ellis, 2003; Ellis & Barkhuizen, 2005; Housen, Kuiken, & Veder, 2012; Skehan, 1998, 2009), many scholars and researchers take a further step to investigate the interaction between the three linguistic dimensions, in specific under the influence of LACM[2] (the Limited Attentional Capacity Model) (Lin & Chen, 2015; Skehan, 1998, 2001, 2003, 2009; Skehan & Foster, 1999, 2001; Yeh & Lin, 2015). A common assumption inherent in LACM in relation to EFL/ESL writing is this: writing under pressure, whether external or internal[3], would greatly consume a writer's limited attentional capacity in brain, causing the CAF to compete with each other, thus forming a trade-off relationship between them as a result (cf. Skehan & Foster, 2001; Housen, Kuiken, & Vedder, 2012).

Interestingly, however, although many researchers continue to devote themselves to this line of inquiry, they have reached no conclusive results. Perhaps one of the major reasons leading to such inconsistence is that these studies try to examine multiple variables at the same time, such as blended variables of time pressure and task complexity. Also, in each study different time frames and writing genres (such as argumentative essays or descriptive writing) are examined (Lin & Chen, 2015). While such approaches are valid in themselves, they somewhat leave the sheer effects of each variable difficult to interpret. Such mutually affecting variables may explain why some scholars would find the development of fluency at the cost of complexity (e.g., Ellis & Yuan, 2004) whereas some others note a different trade-off model: accuracy versus complexity (Skehan, 1996; Skehan & Foster, 1997, 2001). An even stark phenomenon is that some researchers (such as Johnson, Mercado, & Acevedo, 2012) revealed no evidence of any trade-off models in students' CAF even when they wrote in different conditions, a finding that would severely undermine the validity of Skehan's (1998) theory of LACM.

Seeing the conflicts in current literature on CAF, recent researchers (e.g., Yeh & Lin, 2015; Lin & Chen, 2015; Lin, Chen, & Chen, 2015) have suggested more evidence be sought

---

[1] It is commonly suggested that the more fluently or correctly a person writes, the more advanced his/her English writing/speaking ability is (e.g., Fellner & apple, 2006; Lin, 2012).

[2] LACM originates from psycholinguistics and cognitive linguistics, arguing that there is only limited attentional resources in one's brain.

[3] Some examples of external and internal pressure are writing anxiety, motivation towards writing, language users' linguistic abilities, or task complexity. Please see Housen, Kuiken, and Vedder (2012) for more discussion.

to shed greater light on this gap, especially using a stand-alone variable. Among the many variables, time has been suggested to a crucial factor to look at (Lin & Chen, 2015; Yeh & Lin, 2015) mostly because it has been widely cross-examined with other variables in various studies but never on its own. Such lack is even striking in Taiwan's EFL/ESL writing contexts. In an attempt to do so, Lin and Chen (2015) successfully identified a twofold trade-off model in writing samples produced under different time frames by Taiwanese EFL students: a macro model of complexity versus accuracy/fluency and a micro model of accuracy versus fluency (see Appendix A for illustration), the results of which, on the one hand, may synthesize the disagreement on part of previous CAF studies (e.g., Ellis & Yuan, 2004); Skehan & Foster, 1997, 2001) and, on the other, emphasize the need of more investigations so as to identify the exclusive effect of time on CAF. This is specifically necessary given that the findings of Lin and Chen were drawn from a small-scaled case study[4].

An equally important aspect that has not yet been addressed in prior CAF studies in Taiwan is EFL student writers' English proficiency. This being said does not suggest that blends of different variables (time and writing abilities) be investigated here. Rather, the current researcher believes that articles created by writers of different proficiency should not only be studied as a whole to reflect the norm of Taiwanese EFL writers' CAF performance, but such samples should also be treated respectively in accordance with their proficiency levels to portray the fine interactions (if any) of CAF that may differ between writers of different skills (Lin & Chen, 2015), subject to LACM. Such a practice is expressly essential when taking into consideration how the development of student writers' skills change over the course of learning (cf. Hunt, 1965).

Clearly, while writing CAF has received much research attention, more investigation is still needed, in particular in the EFL writing context of Taiwan. For this very reason, the current project was given the aim of filling the gap by carrying out an investigation on a batch of timed writing samples collected from the General English Proficiency Test (GEPT) High Intermediate Level[5]. It is anticipated that the results of this project may well answer the two sets of research questions below:

1-1. Does LACM play a role in the trade-off relationship between the CAF in Taiwanese GEPT High Intermediate test-takers' performance in timed writing?

1-2. If it does, what is the trade-off model like between the CAF?

2-1. Does LACM play a role in the trade-off relationship between the CAF in Taiwanese GEPT High Intermediate test-takers' performance in timed writing when writing

---

[4] Lin and Chen (2015) only investigated 43 student writers' articles.

[5] The GEPT contains five levels of tests, inclusive of elementary, intermediate, high-intermediate, advanced, and superior, with elementary level aligning to A2 Level of the Common European Framework of Reference for Languages (CEFR); intermediate, B1; high-intermediate, B2; advanced; C1; and superior, C2.

qualities are taken into consideration?

2-2. If it does, what are the trade-off models like between the CAF in those writing samples of various qualities?

DEFINITIONS OF CAF

Writing complexity

Writing complexity is arguably the most 'complex' feature to define. In current literature, writing complexity is determined by many to comprise two major aspects: the total number of linguistic units involved and the variety of linguistic patterns used. However, because there are a great deal of various linguistic features in language systems, the measures of the two dimensions seem tedious, if not infeasible. To solve this problem, alternative methods of assessment are suggested. Among the many, the number of T-units, coined by Hunt (1965), is commonly used to indicate complexity levels of textual outputs (e.g., Biber, Gray, & Poonpon, 2011; Ellis & Yuan, 2004; Lin & Chen, 2015; Wolfe-Quintero, Inagaki, & Kim, 1998; Yang, Lu, & Weigle, 2015), mostly in association with the number of sentences or clauses used in the same text (e.g., Ishikawa, 1995; Lin & Chen, 2015; Lu, 2010; Yang, W., Lu, X., & Weigle, S. C. (2015).). Such practices greatly increase the feasibility of assessing writing complexity on the one hand and enhance the validity of measuring complexity on the other, given the strong correlation found between T-unit-related measures and advanced writing samples (Hunt, 1965). For this current study, the researcher thus considered T-units per sentence (T/S) for measuring writing complexity, a practice widely considered (cf. Lin, 2015; Lin & Chen, 2015; Wolfe-Quintero, Inagaki, & Kim, 1998).

Writing accuracy

In contrast to complexity, accuracy is easier to define and measure. As the name suggests, writing accuracy concerns how correct linguistic features are presented. However, as it is difficult to present how correct a feature is, the common measure adopted to show such a quality is the total number of linguistic errors/mistakes found, with more errors/mistakes suggesting lower accuracy, and less errors/mistakes indicating higher accuracy. Instead of directly using the total number of errors, this type of assessment, like that of complexity, is often associated with other important linguistic characteristics, such as T-units, forming typical measures like error-free T-units (EFT)—the total number of T-units identified without any errors (cf. Lin & Chen, 2015; Polio, 1997). As is implied, the greater number of EFT found in an article, the more accurate that article is considered. EFT is thus used for this project (Lin, 2015).

Writing fluency

Although not fully focusing on 'writing' performance, Brown (1994, 2001) offers a

somewhat fuller observation about fluency in language performance, which focuses on the 'flow' of language, or more precisely, a steady flow of language devoid of any correction within a short timeframe. Brown's definition is widely reflected in contemporary measures of fluency in literature, in which "words per minute" is likely to be the most applied assessment (see Ellis and Yuan (2004), Freed (2000), Housen and Kuiken (2009), Ishikawa (1995), Latif (2013), Lin (2015), and Lin and Chen (2015), for example). Although time is considered as an important element in measuring fluency, in this study no precise time lengths spent for each writing sample were recorded, so only the total number of words was used to indicate writing fluency instead (cf. Lin, 2012) since all the writing samples collected were produced under the same time frame.

METHODS

The writing samples

A batch of 150 writing samples, provided by the LTTC and randomly selected from the 2013 GEPT High Intermediate Level writing test responses, was analyzed in this project. In this particular writing test, the test-takers were given 50 minutes to complete two parts: a passage of Chinese-English translation and a guided writing task of 150-180 words. This research aimed to investigate the trade-off relationship between the CAF in the timed writing of GEPT test-takers; therefore, only Part 2 Guided Writing responses were analyzed. The samples comprise 37 responses at Band 2, 53 responses at Band 3, 54 at Band 4 and 6 at Band 5, a score distribution which corresponds to that of an operational GEPT High-Intermediate test. The higher mark suggested more advanced writing skills while the lower mark indicated the contrary, with 4 serving as the threshold: pass or fail. All the scores presented here were the averages of those produced by two anonymous trained raters of the Language Training and Testing Center (LTTC)[6], who used the 6-level holistic writing rubric (Appendix B) designed by the center: from 0 to 6. It should be noted that articles marked 0 are those remain unanswered or contain too little information (less than 40 English words) for their quality to be judged, and those assigned 1 either fail to meet the topic or is impossible to be evaluated due to their ill-content or grammar. For the reasons, none of these two levels were recruited for the analysis.

It must be noted that to answer the research questions proposed above, the 150 writing samples would first be examined as a whole, and then divided into two sub-batches of data for detailed investigation: Sub-Batch 1 (those marked 2 and 3) and Sub-Batch 2 (those marked 4 and 5). The division was decided using the threshold for the writing test, as this indicated a notable difference between the two sub-batches in terms of their writing qualities. This splitting up was anticipated to shed light on the fine differences in the CAF interactions between the writing samples of diverse qualities.

CAF raters and measures

Two raters, different from those of the LTTC, were involved in the CAF assessment, one of whom was the researcher himself and the other was an experienced TESOL teacher. Before the raters started the assessment, they had agreed on the criteria for measuring errors and T-units (Hunt, 1965; Young, 1995), as the two are the fundamental elements to form the measures to assess the complexity and accuracy qualities in this study: T-units per sentence (T/S) for complexity and error-free T-units for accuracy. To confirm the agreement of the two

---

[6] Please visit https://www.lttc.ntu.edu.tw/ for more information about the LTTC.

raters in terms of the two measurements, two correlation tests (Pearson's *r*, two-tailed) were performed. The results indicate statistically significant inter-rater reliability for both items, as statistically significant correlation was found for them both (T-units: $r = .998$, $p = .000$; errors: r .981, $p = .000$), ensuring the validity of the data analysis afterwards. The average of the raters' scores was then used for further analysis. It should be noted that the total number of word tokens (for assessing fluency) and that of sentences were automatically generated by computer using software (WordSmith 5.0), so no inter-rater reliability test was needed for them.

Data analysis

In order to observe whether there was a trade-off model between the CAF of the overall timed writing samples, that of Sub-Batch 1, and that of Sub-Batch 2, both inferential statistics and descriptive statistics were used. First, Pearson's *r* was used to examine the interaction between the CAF; descriptive statistics entailing charts were then used to portray further what the interaction was like (cf. Lin & Chen, 2015).

Results

Interaction between the CAF: Writing samples as a whole

While Table 1 presents the basic information of the writing samples (detailed linguistic data for each sample is presented in Appendix C), Table 2 shows rather complex results: a statistically significant positive correlation between the accuracy and fluency (Pearson's *r* = .585, *p* = .000), a statistically significant yet weak negative correlation between complexity and accuracy (Pearson's *r* = -.189 at *p* = .021), but a non-significant correlation between complexity and fluency (Pearson's *r* = -.066 at *p* = .425). These results suggest that in the GEPT High Intermediate writing samples, lengthier articles are likely to contain more accurate linguistic features. Alternatively, the results also mean that student writers who can write more accurately may write relatively faster or more. In contrast, such accuracy may be developed at the cost of complexity, or vice versa, as a statistically significant negative correlation was found here. However, no significant interaction was found between writing fluency and complexity.

Table 1. Basic information of the writing samples

Items	N	Minimum	Maximum	Mean	SD
Sentences	150	3.00	25.00	12.27	4.44
T-units	150	4.00	26.00	13.60	4.58
Errors	150	.50	45.50	15.95	9.96
T/S (Complexity)	150	.85	3.00	1.14	.25
EFT (Accuracy)	150	.00	15.50	5.54	3.81
Tokens (Fluency)	150	61.00	353.00	205.61	62.39

Table 2. Correlation tests (Pearson's *r*) for CAF interactions

	Complexity	Accuracy	Fluency
Complexity	1		
Accuracy	-.189*	1	
Fluency	-.066	.585***	1

Descriptive statistics: interactions between the CAF

Figure 1 provides further evidence that LACM has an effect in the participants' writing samples. To be specific, despite of some exceptions, the *complexity* curves remained similar in shape, slope, and height, complementing the inferential finding in which very small negative correlation coefficient was found between complexity and the others. However, both *fluency*

and *accuracy* curves changed drastically and seemingly raised together and dropped together, suggesting a corresponding interaction between each other. This further verifies the positive correlation found between accuracy and fluency as shown above.
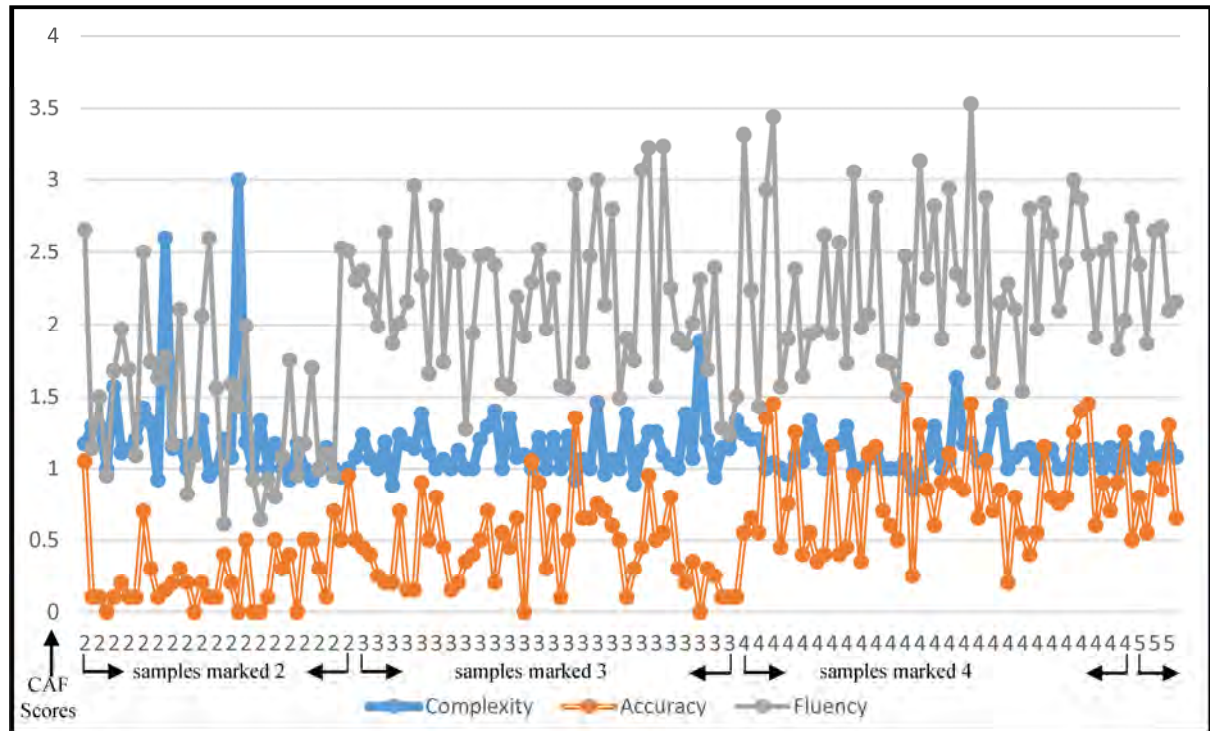

Figure 1. The CAF interactions in the overall writing samples
(*Note*. The dots in the figure represent the scores for each linguistic feature.)

Interaction between the CAF: Sub-Batch 1

While Table 3 presents the basic information of the writing samples of Sub-Batch 1, Table 4 shows a statistically significant positive correlation between the accuracy and fluency (Pearson's $r = .517$, $p = .000$) but no significant correlation was found between complexity and accuracy (Pearson's $r = -.197$ at $p = .063$) or between complexity and fluency (Pearson's $r = -.017$ at $p = .874$). Mostly resembling those of overall writing samples, these results suggest that despite their relatively poor writing quality, lengthier articles of this batch still tend to contain more accurate linguistic features, or that student writers who can write more accurately are likely to develop their fluency as well. Although a clear mutual benefit can be seen between fluency and accuracy, no clear evidence shows the way by which complexity development entered this interrelationship.

Descriptive statistics: interactions between the CAF in Sub-Batch 1

Figure 2 provides further evidence that LACM has an effect on test-takers' writing performance. Clearly, despite of some exceptions, the shape and slope of the *complexity* curve remained mostly the same from right to left, which may again be taken as a sign to support the inferential finding that no correlation was there between complexity and the other two features. However, both *fluency* and *accuracy* curves changed drastically and seemingly raised together and dropped together, implying a corresponding interaction between each other. This further verifies the positive correlation found between accuracy and fluency as shown above, in spite of the relatively low writing quality of Sub-Batch 1.

Table 3. Basic information of Sub-Batch 1

Items	N	Minimum	Maximum	Mean	SD
Sentences	90	3.00	24.00	11.59	4.58
T-units	90	4.00	25.50	13.04	4.80
Errors	90	2.00	45.50	19.46	9.53
T/S (Complexity)	90	.88	3.00	1.17	.30
EFT (Accuracy)	90	.00	13.50	3.77	2.92
Tokens (Fluency)	90	61.00	324.00	187.96	63.03

Table 4. Correlation tests (Pearson's *r*) for CAF interactions in Sub-Batch 1

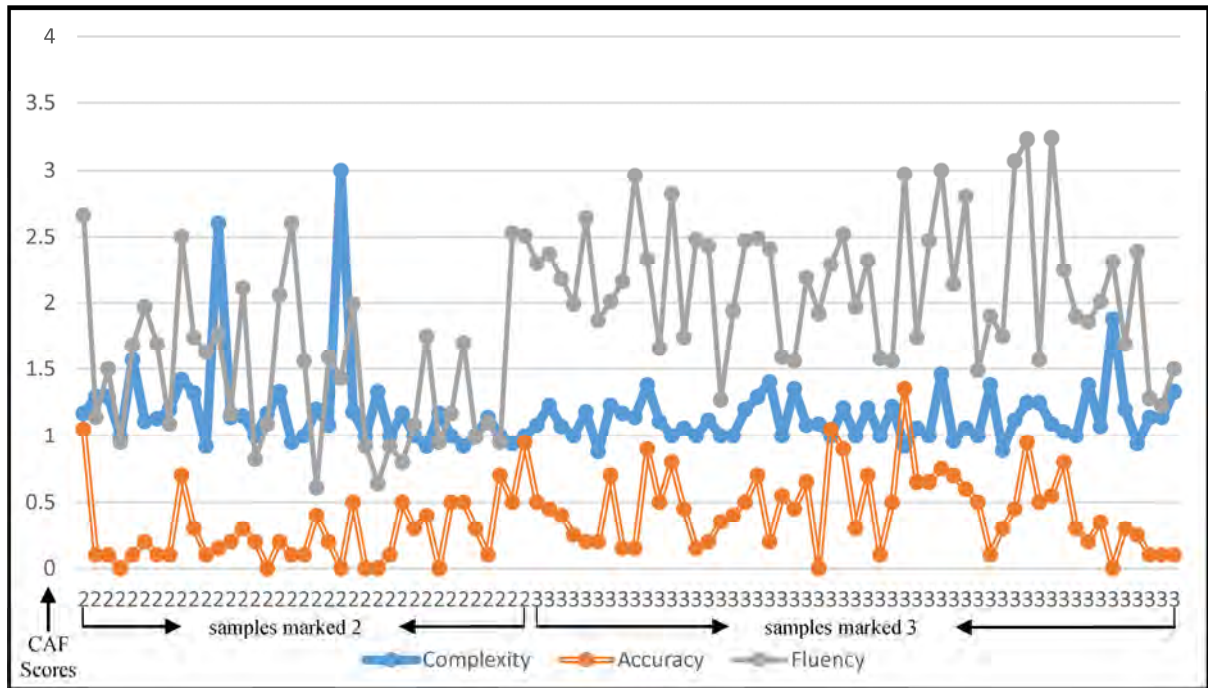|	Complexity	Accuracy	Fluency
Complexity	1		
Accuracy	-.197	1	
Fluency	-.017	.517***	1

Figure 2. The CAF interactions in the Sub-Batch 1

(*Note.* The dots in the figure represent the scores for each linguistic feature.)

Interaction between the CAF: Sub-Batch 2

Table 5 presents the basic information of Sub-Batch 2; Table 6 shows the inferential results, in which a statistically significant positive correlation was found between the accuracy and fluency (Pearson's $r = .501$, $p = .000$). However, no statistically significant correlation was detected between complexity and accuracy (Pearson's $r = -.020$ at $p = .882$) or between complexity and fluency (Pearson's $r = -.047$ at $p = .758$). These results also suggest that in the advanced GEPT High Intermediate writing samples, lengthier articles are more likely to contain more correct grammar, or that skilful student writers who can write more correctly may also write faster or more in the given time. In contrast, however, the development of either accuracy or fluency had no obvious interrelationship with that of complexity.

Table 5. Basic information of Sub-Batch 2

Items	N	Minimum	Maximum	Mean	SD
Sentences	60	5.00	25.00	13.28	4.06
T-units	60	6.00	26.00	14.43	4.13
Errors	60	.50	41.00	10.68	8.16
T/S (Complexity)	60	.85	1.63	1.10	.13
EFT (Accuracy)	60	2.00	15.50	8.19	3.45

| Tokens (Fluency) | 60 | 143.00 | 353.00 | 232.07 | 51.49 |

Table 6. Correlation tests (Pearson's *r*) for CAF interactions in Sub-Batch 2

	Complexity	Accuracy	Fluency
Complexity	1		
Accuracy	-.020	1	
Fluency	-.041	.501***	1

Descriptive statistics: interactions between the CAF in Sub-Batch 2

Figure 3 also shows a clear interaction between accuracy and fluency, which may suggest that LACM has an effect on test-takers' writing performance. This is evidenced by how both *fluency* and *accuracy* curves changed drastically and seemingly raised together and dropped together. However, like those in Figures 1 and 2, the shape and slope of the *complexity* curve remained mostly flat from right to left, which may again serve as a sign of supporting the inferential finding that no correlation was there between complexity and the other two features. Clearly, the relatively high writing quality of Sub-Batch 2 contained no different CAF interaction from that shown in low quality writing.
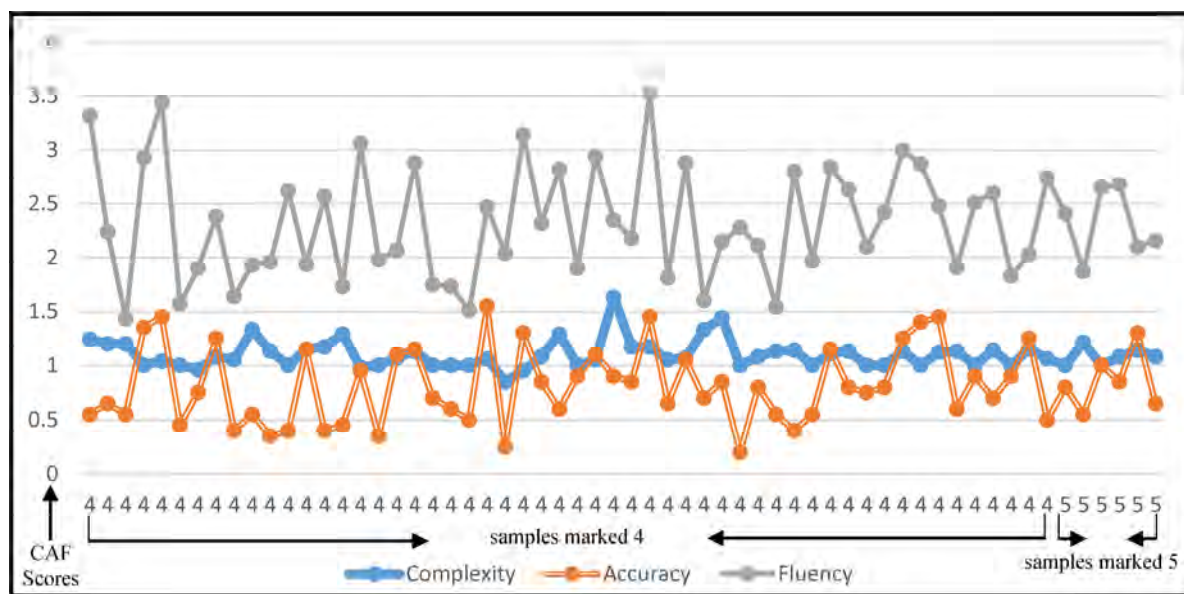

Figure 3. The CAF interactions in the Sub-Batch 2
(*Note*. The dots in the figure represent the scores for each linguistic feature.)

Discussion and conclusions

The present study investigated whether or not LACM might play a role in the trade-off relationship between the CAF in Taiwanese GEPT High Intermediate test-takers' writing performance when writing samples of various qualities were investigated as a whole. The data were first examined as a whole and then respectively according to their writing qualities (two batches of different levels). This was done in the hope of portraying not only the interaction between the CAF subject to LACM, but also the fine interrelationship in CAF attributed to writing proficiency as an influential factor. Not completely in line with prior results, the results of this study indicate a potential trade-off model; that is, student writers, regardless of proficiency levels, seem to focus on the development of both accuracy and fluency, likely at the cost of complexity, an interrelationship meriting discussion.

To begin with, although no statistically significant correlations were found between complexity and fluency/accuracy, the sharp contrast between the consistent *negative* interrelationships of complexity versus fluency/accuracy and the *positive* interrelationship of fluency and accuracy has highlighted this possibility: accuracy and fluency are likely to be developed at the cost of complexity. The reason why such correlations was not fully evidenced by statistical significance may be subject to the sample sizes used for this project. Such interpretation seems probable when considering only in the overall samples (150 articles) was the complexity found in a significant (although weak) negative correlation with accuracy, but never in Sub-Batches 1 (90 articles) and 2 (60 articles). The decreasing possible correlations between fluency and accuracy from the data as a whole ($r = .587$) through Sub-Batches 1 (.517) and 2 (.501) lend further support to this reasoning. Given this, future researchers may consider relatively large samples to re-test the trade-off model claimed here in this study.

Second, confirming the LACM theory, the results also echo that by Skehan (1996) and Skehan and Foster (1997, 2001), in that accuracy competes with complexity and also that by Ellis and Yuan (2004), who finds fluency development at the expense of complexity (e.g., Ellis & Yuan, 2004). Such interpretation may seem in conflict, but to some extent, like what Lin and Chen (2015) finds, the results of this study may actually synthesize those by Skehan (1996), Skehan and Foster (1997, 2001), and Ellis and Yuan (2004), supporting the part of the formulation drawn by Lin and Chen (2015): a macro trade-off model (complexity versus accuracy/fluency). In contrast to Lin and Chen, however, while such a model entails a micro interrelationship (accuracy versus fluency), the micro interrelationship found in this study is mutually beneficial to each other.

One possible reason for the formulation of this study to differ in part from that by Lin and Chen (2015) in terms of the micro model is likely ascribed to the different levels of time pressures imposed on the EFL writers. While in the current study all the writers were given

the same amount of time to write, in Lin and Chen's study the EFL student writers were required to compose different articles under different time frames. In other words, in addition to examining timed writing, Lin and Chen also looked at the influence of different time pressures on Taiwanese EFL student writers' performance, in which greater time pressure clearly affects the balance between accuracy and fluency, causing them to compete with each other. Given this, it may seem valid at this point to conclude that not only does the current study ensure there is a micro model between fluency and accuracy, but it indicates time as one of the most influential factors affecting its balance, in turn justifying the need of investigating time as a sheer variable in studies of this type. This interpretation is further strengthened by this: the results of the current study also echoes the early study by Kuiken and Vedder (2008), who report that language proficiency levels have no effects in the interrelationship between complexity and accuracy.

In addition to contributing to the understanding of LACM in timed writing produced by GEPT High Intermediate test-takers, the research focus and methods of this investigation also open up a few opportunities for further studies although the design of this study is valid itself. First of all, while time frames were not considered in assessing fluency in this study, future researchers may take this element into fluency measurement as in many previous investigations (cf. Ellis & Yuan, 2004; Freed, 2000; Housen & Kuiken, 2009; Ishikawa, 1995; Latif, 2013; Lin & Chen, 2015), the practice of which may shed greater light on the fine interaction between CAF, precisely between fluency and accuracy given their positive interaction. Also, more investigations may be done to examine whether similar patterns found in this study are repeated in the CAF interaction happening in the writing samples of the other GEPT levels, such as advanced or elementary. Following this line of inquiry shall portray the CAF trade-off models to an even fuller extent. In a similar note, an interesting topic to develop further is a closer look at the fine interaction between complexity and accuracy/fluency. While this study temporarily proved the existence of such relationship, it remains unsolved what, or more specifically which linguistic feature in complexity, is traded off when test-takers focus on the accuracy/fluency. To complement the results of this study or any future studies as suggested here, researchers in this field may consider qualitative inquiries, a practice that is rarely seen in this field. In addition to depicting the exact CAF interaction in Taiwanese GEPT test-takers' writing samples, it would be helpful to understand exactly why they would choose to develop one (or two) linguistic dimension(s) over another. In a related note, future researchers may create new stances on observing how CAF may change differently between writing samples of various qualities, an idea inspired by the result that in this study the writing quality of Sub-Batch 1 was found slightly more complex than that of Sub-Batch 2 (see Tables 3 and 5)—a phenomena seemingly contradicting the general assumption that advanced writers would write with more complexity than less skillful writers. As this aspect does meet the research goal of this study, no sufficient data have been

generated to cope with it, leaving some room for future studies to fill in. Last but not least, future researchers may extend from the understanding of CAF interactions to examining how such interrelationships may serve as predictors of students' writing scores. It would be helpful to perform this task using logistic regression. Doing this may bring about different perspectives on the meaning of CAF interactions, in turn justifying the need for more CAF studies in future.

Acknowledgements

References

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*(1), 59-84.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*(1), 59-84.

Evans, N. W., Hartshorn, K. J., Cox, T. L., & Martin de Jel, T. (2014). Measuring written linguistic accuracy with weighted clause ratios: A question of validity. Journal of Second Language Writing, 24, 33-50.

Groom, N., & Littlemore, J. (2011). Doing applied linguistics: A guide for students. London and New York: Routledge.

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency. In A. Housen, F. Kuiken, & I. Vedder (Eds.), Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA (Vol. 32) (pp. 1-20). Amsterdam: John Benjamins Publishing Company.

Ishikawa, S. (1995). Objective measurement of low-proficiency EFL narrative writing. Journal of Second Language Writing, 4, 51-70.

Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. Journal of Second Language Writing, 21, 264-282.

Latif, M. M. M. A. (2013). What do we mean by writing fluency and how can it be validly measured? Applied Linguistics, 34(1), 99-10.

Lin, M. H. (2012). Blog assisted language learning in the EFL writing classroom: An experimental study. Unpublished Doctoral dissertation. University of Birmingham, Birmingham, UK.

Lin, M. H. (2014). Effects of Classroom Blogging on ESL Student Writers: An Empirical Reassessment. The Asia-Pacific Education Researcher, 23(3), 577-590. DOI: 10.1007/s40299-013-0131-8.

Lin, M. H. (2015). Writing complexity, accuracy, and fluency: Validity as indicators for Taiwanese EFL writers? Paper presented at The 1st Interschool Academic Conference on Linguistics, Literature, and Language Teaching, 2015. Providence University,

Taichung, Taiwan.

Lin, M. H., & Chen, I.-T. (2015). Time factors in writing complexity, accuracy, and fluency: A preliminary trade-off model found in Taiwanese EFL students' compositions. In the Proceedings of 2015 International Conference and Workshop on TEFL & Applied Linguistics (pp. 161-169). Taipei: Crane Publishing.

Lin, M. H., Chen, I.-T., & Chen, H.-K. (June, 2015). What does time buy? A preliminary investigation on Taiwanese EFL students' writing performance. Paper presented at the 11th Annual English Conference. New Taipei City, Taiwan.

Lin, M. H., Li, J. J., Hung, P. Y., & Huang, H. W. (2014) Blogging a Journal: Changing Students' Writing Skillsand Perceptions. ELT Journal, 68(4), 422-431. DOI: 10.1093/elt/ccu032

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. International Journal of Corpus Linguistics, 15(4), 474-496.

Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. Journal of Second Language Writing, 19(4), 218-233.

Polat, B., & Kim, Y. (2013). Dynamics of Complexity and Accuracy: A Longitudinal Case Study of Advanced Untutored Development. Applied Linguistics. DOI: 10.1093/applin/amt013

Polio, C. G. (1997). Measures of linguistic accuracy in second language writing. Language Learning, 47, 101-143.

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. IRAL-International Review of Applied Linguistics in Language Teaching, 45(3), 161-176.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain (Eds.). Researching pedagogic tasks: Second language learning, teaching and testing (pp. 167-185). Harlow: Pearson Education.

Skehan, P. (2003). Task-based instruction. Language Teaching, 36(1), 1-14.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. Applied Linguistics, 30(4), 510-532.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. Language Teaching Research, 1, 185–211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retelling. Language Learning, 49(1), 93-100.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), Cognition and second language instruction, 183-205.

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. Journal of Second Language Writing, 28, 53-67.

Yeh, S.-Y., & Lin, M. H. (May, 2015). The trade-off relationship between English writing fluency, accuracy, and complexity: The time factor in college students' English writing performance. Paper presented at the 32nd International Conference on English Teaching & Learning, Taipei, Taiwan.

Young, R. (1995). Conversational styles in language proficiency interviews. Language Learning, 45(1), 3-42.

Vardazaryan, K. (2012). Second language development and blogging: The influence of blogging on EFL Learners' writing fluency, accuracy and complexity. Deutschland, Germany: Lap Lambert Academic Publishing.

The Author

Dr. Ming Huei Lin is an assistant professor in the English department at Tamkang University, Taiwan. He has wide-ranging interests in TESOL, applied linguistics, and educational technology. He has published articles in journals, including *British Journal of Educational Technology*, *Educational Technology & Society*, *ELT Journal*, *TESOL Quarterly*, and *The Asia-Pacific Education Researcher*. Email: johnlinminghuei@gmail.com

APPENDIX A

A two-folded trade-off relationship between CAF
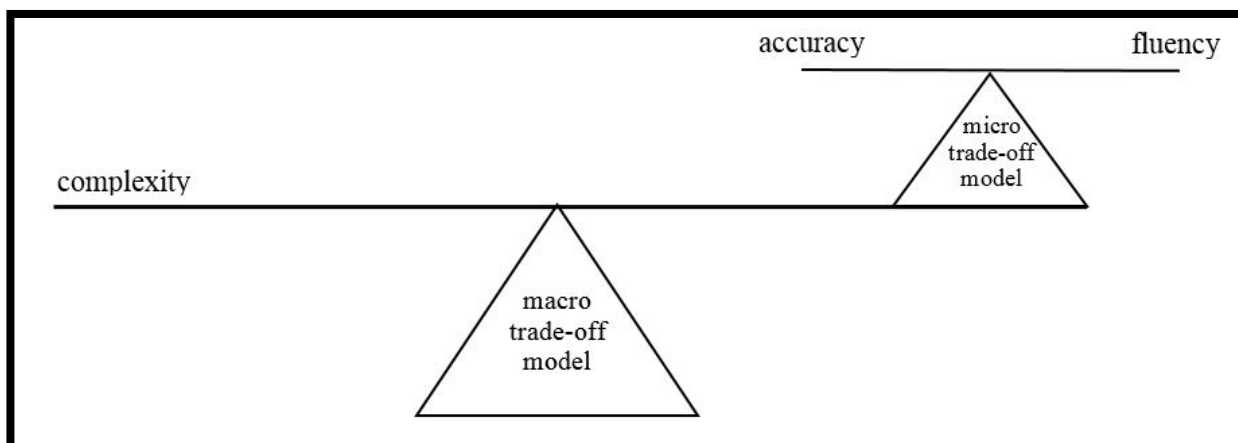**(Figures adopted from Lin and Chen, 2015, p. 166-167)**

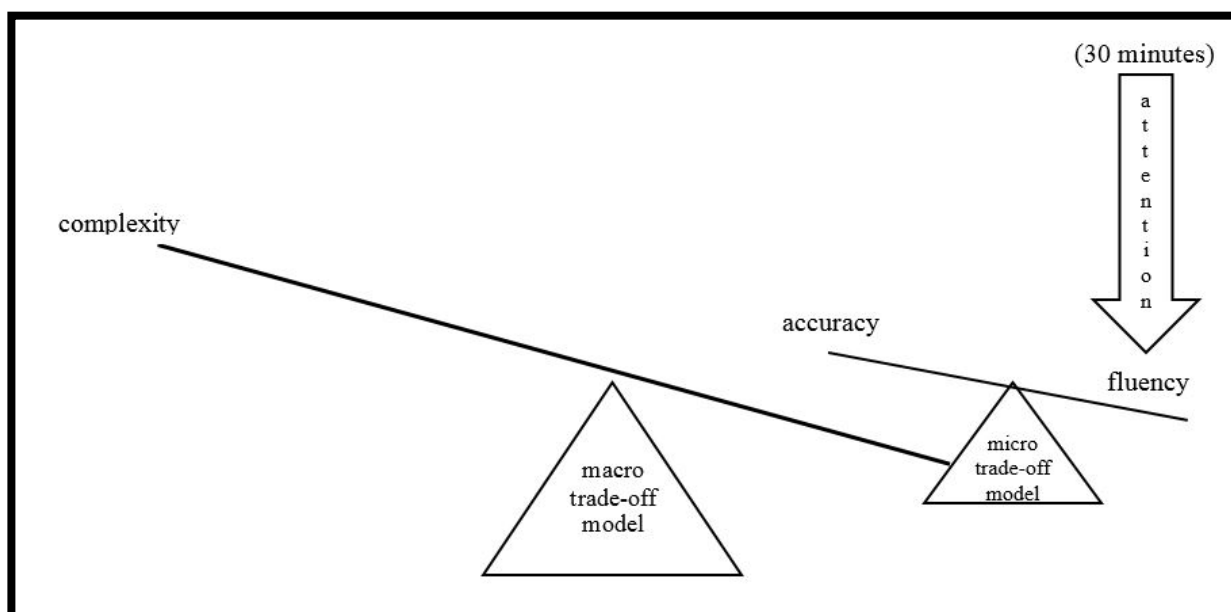Figure A.1. The balanced trade-off model between CAF

Figure A.2. The imbalance between CAF when attention is paid to the development of fluency
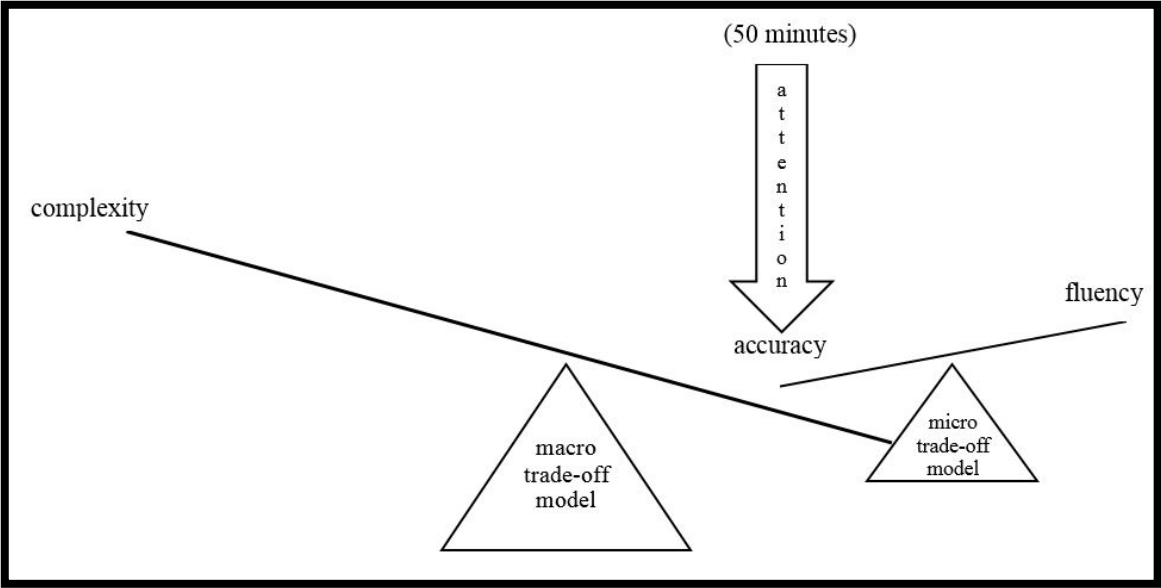

Figure A.3. The imbalance between CAF when attention is paid to the development of accuracy

APPENDIX B

The holistic writing rubric for GEPT High Intermediate Level

Note. The following English version is a translation by the researcher of this project. The original version of this rubric is in Chinese, which can be seen online here: https://www.gept.org.tw/Exam_Intro/t03_introduction.asp

Level	
5	The article expressively and properly addresses the topic; its organization is complete and cohesive. The article also demonstrates a variety of proper vocabulary and sentence patterns. Only few mistakes/errors can be found.
4	The article meets the topic, is well-organized, and has good coherence in general. The article contains mostly accurate vocabulary and sentence patterns; only a few mistakes/errors are found.
3	The article meets the topic in general; its organization is fine but its coherence needs improving. The article contains high-frequent vocabulary and basic sentence structures, but mistakes/errors are often found in advanced vocabulary or complex sentence patterns.
2	The article only meets part of the topic; its organization is incomplete and lacks coherence; its vocabulary is limited; and it contains mostly basic sentence structures, in which many mistakes/errors are found, influencing readers' understanding.
1	The article fails to meet the topic, is ill-organized, shows rather limited vocabulary, and contains a variety of problematic sentence structures severely preventing readers' understanding.
0	Unanswered

APPENDIX C

Detailed linguistic information for each writing sample

Students	Scores	Tokens	Sentences per T-unit	Error Free T-units
1.	2.00	266.00	1.17	10.50
2.	2.00	114.00	1.29	1.00
3.	2.00	150.00	1.29	1.00
4.	2.00	95.00	1.00	.00
5.	2.00	168.00	1.57	1.00
6.	2.00	197.00	1.11	2.00
7.	2.00	169.00	1.13	1.00
8.	2.00	109.00	1.20	1.00
9.	2.00	250.00	1.42	7.00
10.	2.00	174.00	1.32	3.00
11.	2.00	163.00	.92	1.00
12.	2.00	177.00	2.60	1.50
13.	2.00	117.00	1.14	2.00
14.	2.00	211.00	1.15	3.00
15.	2.00	82.00	1.00	2.00
16.	2.00	109.00	1.17	.00
17.	2.00	206.00	1.33	2.00
18.	2.00	260.00	.95	1.00
19.	2.00	156.00	1.00	1.00
20.	2.00	61.00	1.20	4.00
21.	2.00	159.00	1.08	2.00
22.	2.00	143.00	3.00	.00
23.	2.00	199.00	1.18	5.00
24.	2.00	92.00	1.00	.00
25.	2.00	64.00	1.33	.00
26.	2.00	92.00	1.00	1.00
27.	2.00	80.00	1.17	5.00
28.	2.00	108.00	1.00	3.00
29.	2.00	175.00	.92	4.00
30.	2.00	95.00	1.17	.00
31.	2.00	117.00	1.00	5.00
32.	2.00	170.00	.92	5.00

33.	2.00	99.00	1.00	3.00
34.	2.00	110.00	1.14	1.00
35.	2.00	95.00	1.00	7.00
36.	2.00	253.00	.94	5.00
37.	2.00	251.00	1.00	9.50
38.	3.00	230.00	1.08	5.00
39.	3.00	237.00	1.23	4.50
40.	3.00	218.00	1.07	4.00
41.	3.00	199.00	1.00	2.50
42.	3.00	264.00	1.18	2.00
43.	3.00	187.00	.88	2.00
44.	3.00	201.00	1.23	7.00
45.	3.00	216.00	1.17	1.50
46.	3.00	296.00	1.14	1.50
47.	3.00	233.00	1.38	9.00
48.	3.00	166.00	1.11	5.00
49.	3.00	282.00	1.00	8.00
50.	3.00	174.00	1.06	4.50
51.	3.00	248.00	1.00	1.50
52.	3.00	243.00	1.12	2.00
53.	3.00	127.00	1.00	3.50
54.	3.00	194.00	1.00	4.00
55.	3.00	247.00	1.20	5.00
56.	3.00	249.00	1.29	7.00
57.	3.00	241.00	1.40	2.00
58.	3.00	159.00	1.00	5.50
59.	3.00	156.00	1.35	4.50
60.	3.00	219.00	1.08	6.50
61.	3.00	192.00	1.09	.00
62.	3.00	229.00	1.00	10.50
63.	3.00	252.00	1.21	9.00
64.	3.00	197.00	1.00	3.00
65.	3.00	232.00	1.21	7.00
66.	3.00	158.00	1.00	1.00
67.	3.00	156.00	1.22	5.00
68.	3.00	297.00	.92	13.50
69.	3.00	174.00	1.06	6.50

70.	3.00	247.00	1.00	6.50
71.	3.00	300.00	1.46	7.50
72.	3.00	214.00	.96	7.00
73.	3.00	280.00	1.06	6.00
74.	3.00	149.00	1.00	5.00
75.	3.00	190.00	1.38	1.00
76.	3.00	175.00	.89	3.00
77.	3.00	307.00	1.12	4.50
78.	3.00	323.00	1.25	9.50
79.	3.00	157.00	1.25	5.00
80.	3.00	324.00	1.09	5.50
81.	3.00	225.00	1.03	8.00
82.	3.00	190.00	1.00	3.00
83.	3.00	186.00	1.38	2.00
84.	3.00	201.00	1.07	3.50
85.	3.00	231.00	1.88	.00
86.	3.00	169.00	1.20	3.00
87.	3.00	239.00	.94	2.50
88.	3.00	128.00	1.14	1.00
89.	3.00	123.00	1.14	1.00
90.	3.00	150.00	1.33	1.00
91.	4.00	332.00	1.24	5.50
92.	4.00	224.00	1.20	6.50
93.	4.00	143.00	1.20	5.50
94.	4.00	293.00	1.00	13.50
95.	4.00	344.00	1.04	14.50
96.	4.00	157.00	1.00	4.50
97.	4.00	190.00	.96	7.50
98.	4.00	238.00	1.08	12.50
99.	4.00	164.00	1.05	4.00
100.	4.00	193.00	1.33	5.50
101.	4.00	196.00	1.13	3.50
102.	4.00	262.00	1.00	4.00
103.	4.00	194.00	1.15	11.50
104.	4.00	257.00	1.17	4.00
105.	4.00	173.00	1.29	4.50
106.	4.00	306.00	.98	9.50

107.	4.00	198.00	1.00	3.50
108.	4.00	207.00	1.07	11.00
109.	4.00	288.00	1.13	11.50
110.	4.00	175.00	1.00	7.00
111.	4.00	173.00	1.00	6.00
112.	4.00	151.00	1.00	5.00
113.	4.00	247.00	1.06	15.50
114.	4.00	204.00	.85	2.50
115.	4.00	314.00	.95	13.00
116.	4.00	232.00	1.08	8.50
117.	4.00	282.00	1.29	6.00
118.	4.00	190.00	1.00	9.00
119.	4.00	294.00	1.05	11.00
120.	4.00	235.00	1.63	9.00
121.	4.00	218.00	1.17	8.50
122.	4.00	353.00	1.17	14.50
123.	4.00	181.00	1.05	6.50
124.	4.00	288.00	1.07	10.50
125.	4.00	160.00	1.33	7.00
126.	4.00	215.00	1.44	8.50
127.	4.00	228.00	1.00	2.00
128.	4.00	211.00	1.08	8.00
129.	4.00	154.00	1.13	5.50
130.	4.00	280.00	1.14	4.00
131.	4.00	197.00	1.00	5.50
132.	4.00	284.00	1.11	11.50
133.	4.00	263.00	1.13	8.00
134.	4.00	210.00	1.00	7.50
135.	4.00	242.00	1.00	8.00
136.	4.00	300.00	1.13	12.50
137.	4.00	287.00	1.00	14.00
138.	4.00	248.00	1.12	14.50
139.	4.00	191.00	1.13	6.00
140.	4.00	251.00	1.00	9.00
141.	4.00	260.00	1.14	7.00
142.	4.00	183.00	1.00	9.00
143.	4.00	203.00	1.15	12.50

144.	4.00	274.00	1.06	5.00
145.	5.00	241.00	1.00	8.00
146.	5.00	187.00	1.21	5.50
147.	5.00	265.00	1.00	10.00
148.	5.00	268.00	1.08	8.50
149.	5.00	210.00	1.14	13.00
150.	5.00	216.00	1.08	6.50