

The Accuracy of Students' Self-Assessment-How Useful Is a Checklist in EAP Writing

Longdi Chen¹ & Jie Zhang¹

¹ School of Languages, Xi'an Jiaotong Liverpool University, Suzhou, China

Correspondence: Longdi Chen, School of Languages, Xi'an Jiaotong Liverpool University, Suzhou, China.
E-mail: Longdi.Chen@xjtlu.edu.cn

Received: February 2, 2023

Accepted: February 27, 2023

Online Published: February 28, 2023

doi: 10.5539/elt.v16n3p23

URL: <https://doi.org/10.5539/elt.v16n3p23>

Abstract

Checklists, as a form of self-assessment, have been provided to students by educators as a way to promote learner autonomy. However, due attention has not been paid to the accuracy of those completed checklists. Through a survey, interviews, and comparison of students' and researchers' completed checklists, the research intends to investigate the accuracy of students' self-assessment of their writing in a Sino-foreign university in China. Suggestions on checklist design and pedagogical implications on how to improve learners' self-assessment are made.

Keywords: self-assessment, accuracy, writing, EAP

1. Introduction

Learner autonomy has long been valued especially against the background of globalization where institutions and organizations cultivate global talents. As an important part that constitutes learner autonomy, students' self-assessment of their learning and performance seems to be more challenging and of profound meaning. According to Brown and Harris (2013), self-assessment (SA) is a technique that students can use to evaluate and compare their learning with expected learning outcomes and it serves many desirable functions in teaching and learning. Andrade (2019) asserted that the purpose of SA was to produce feedback that in return improves students' performance and consolidates learning. In fact, in practice, SA has been adopted by the Common European Framework of Reference for Languages (CEFR), as an instrument to capture and understand language performance (Liu & Brantmeier, 2019). It has been widely used in various scenarios: for self-placement, for achievement (before and after self-assessing questionnaires), and for particular skills and language components assessments (Bailey, 2003).

According to Andrade (2019), different activities fall into the category of SA, from simply assigning a happy or sad face to using rubrics to evaluate one's writing and reflective journal writing. She particularly mentioned the research by Panadero, Alonso-Tapia, & Huertas (2014) and the list of SA criteria used in the study and referred to this particular tool (scripts) as "checklist" in classroom assessment context. However, since such instruments are used by students for SA, teachers seldom check the completion or accuracy of these materials; therefore, there has been a lack of understanding of the accuracy of students' SA. This will largely undermine the function and purpose of students' SA as teachers would not be able to address students' identified weaknesses without proper investigation and analysis of students' SA. In her critical review of SA, Andrade (2019) declared that SA would be meaningless if opportunities for correction and adjustments were not afforded. This is particularly a problem for teachers with Chinese students who are more accustomed to the authoritative teaching style where teachers evaluate and manage most of students' studies (Cheng and Ding, 2021). How they would assess their own work, how accurate they would be, and more importantly, how we as teachers can assist them and improve their SA accuracy will be of great importance to investigate. The current study intends to gain an insight and provide some suggestions in this regard.

2. Literature Review

Self-assessment (SA), as a kind of formative assessment, refers to "the procedure by which learners themselves evaluate their language skills and knowledge" (Bailey, 2003, p227). Learners are required to activate self-monitoring, examine, and reflect on their own work. During the process, learners need to decide whether

their work has met the criteria and if revisions are necessary (Andrade and Valtcheva, 2009). Zimmerman and Schunk (2001) believe that learners are becoming active behaviourally, metacognitively, and motivationally in this process. This is clearly indicated in the term itself and can be easily understood. As the name suggests, students have to be engaged in the assessment process themselves and this means they have to activate their awareness and monitoring mechanism to participate and become autonomously driven. In terms of the motivational boost, Butler & Lee (2010) attribute the benefit to a relatively low-anxiety setting SA creates. Other benefits of SA include: it promotes learner autonomy (Dann, 2002), facilitates learners' understanding of the criteria of the assessments (Lewis, 1990), eases teacher burden (Butler & Lee, 2010), facilitates self-regulated learning (Panadero, Jonsson and Botella, 2017; Butler & Lee, 2010) and may even have long-lasting effects on learners' future learning and language development (Butler & Lee, 2010).

As a natural and integral part of language learning, it is reasonable that SA has long been valued and encouraged by language instructors. However, one significant drawback of SA is the subjectivity and inaccuracy of learners' judgment (Bailey, 2003) and Bulter and Lee (2010) considered the inherent subjectivity of SA as a threat to the validity of SA. Current research has attempted to establish the correlation of SA accuracy with different factors. For example, both Ross (2006) and Eva, Cunnington, Reiter, Keane, & Norman (2004) mentioned that under-achieving students tend to overestimate their performance. Brantmeier, Vanderplank, & Strube (2012) investigated students' SA accuracy and their proficiency level and found that language proficiency level was positively correlated with SA accuracy. According to Ashton (2014), the positive correlations between SA of language abilities have been revealed by many empirical studies although Ross (2006) also mentioned a common observation that some good students may under-evaluate their attainment, which indicated the influences of other factors. Liu and Brantmeier (2019) summarized different factors that may influence learners' SA accuracy: SA constructs (e.g., criterion-referenced or not, contextualized or decontextualized), SA experience and training, target language proficiency, receptive or productive language skills, and so forth.

Among the different factors, another commonly researched SA aspect is criteria-referenced SA. According to Andrade (2000), a rubric is a document of a list of criteria and it can be used as a SA instrument (Stiggins, 2001). Andrade, Mycek, & Du (2010) engaged 162 middle school students and have them self-assess their written assignments with a rubric. It was found that reading a model, generating criteria, and using a rubric contributed to more effective writing among these students. Similarly, Kim (2019) investigated SA in essay writing of Korean high school students and found that rubric-referenced SA positively affected students writing. Andrade (2019) reviewed her studies on SA and admitted that she had favored SA regarding standards such as rubrics and checklists and Liu and Brantmeier (2019) stated that it is generally believed that criteria-referenced SA is more positively correlated to SA accuracy.

In the focal university where the study was conducted, we also assess students through criteria-referenced assessments. For example, for most of the writing courseworks, students are assessed using the marking descriptor which clearly states the criterion for each band. As the descriptor itself can be challenging for students to interpret, as language lecturers, we usually provide students with a checklist, which is often designed to mimic (or at least cover different aspects of) the marking descriptor to facilitate SA. Therefore, the checklist is a quasi-criteria-referenced SA tool in this case. However, due attention has not been paid to those completed checklists to examine exactly how accurate students' SA is. It is especially crucial considering that Chinese students, the majority of the focal university student body, have been primarily assessed by standardized tests (Carless, 2005). In fact, Bulter and Lee (2010) declared that implementation of SA may be challenging where students study in a highly competitive environment and mainly receive teacher-centered instruction. Therefore, it would be intriguing to investigate the accuracy of their SA. Hence, the research intends to address the following research questions:

- (1) How accurate is students' self-assessment using a checklist to assess their EAP writing in the focal university?
- (2) What factors would influence students' accuracy of self-assessment by using a checklist?

3. Research Design and Methodology

3.1 Participants

65 Year 2 students from both investigators' EAP105 (an undergraduate EAP module for Year 2 business students) teaching groups in a Sino- foreign EMI university in China participated in the research, among whom 6 students also participated in the interview on a voluntary basis. They were chosen because they had a similar educational background and share a similar amount of training and experience of using a checklist to evaluate their own work. All participants produced the first 2 sections of their writing coursework: the introduction of a

business evaluation report and a business proposition. Before their writing, all participants were taught how to complete the task, such as a detailed breakdown of sections, required elements in both sections, paragraph structure, and writing techniques, etc.. After they finished writing the two sections, the participants evaluated their own writing using a checklist (students had to tick YES or NO for each item on the checklist) provided by the researchers and then reflected on their experience of using the checklist by completing a survey containing 4 multiple choice questions and 4 opened ended questions. Researchers collected all the writings, checklists, survey sheets, and consent forms, coded each participant with a number from 1 to 57, and organized the documents accordingly.

3.2 Research Methods and Procedure

The research adopted a mixed approach and the following methods were used: a cross-examination of teacher and student checklist, a questionnaire, and face-to-face interviews. It needs to be pointed out that the items on the checklist reflect the criteria students are assessed by and we integrated designs of similar checklists used by different EAP modules in the focal university and incorporated their strengths. The reason why a cross-examination of teacher and student checklist was conducted was that it is usually a common research design to compare students' assessment with external assessment (in this case, teachers' assessment) and the research design improves the criterion-related validity (Bailey, 2003). More importantly, it would help reveal the differences between teacher and student assessment. In addition, the questionnaire and interview were used to generate more detailed and specific data for interpretation.

The researchers followed the procedures below: 1. The researchers collected participants' writing, checklists, and survey sheets. 2. Both researchers read one participant's writing and used the checklist (the same one students used) to evaluate the participant's performance. 3. Researchers compared their results of checklists. If the results were the same, the previous procedures were repeated; otherwise, comparisons and discussions were conducted and a finalized teacher checklist was produced. 4. The researchers compared how similar students' checklists were to the researchers'. For every item that was the same, one score was awarded and a final score was calculated for each student. 5. Investigators recorded the scoring details of each participant. 6. Investigators recorded students' survey results. 7. Investigators conducted individual interviews (audio recorded) about participants' experience of using the checklist.

3.3 Data Analysis

Data of the current study comes from 65 survey results and 57 scripts and teacher and student checklists and 6 interviews. All survey results were computed and transferred to Excel for analysis and graphs were generated to better understand the percentage and distribution of different choices. Each student's scoring of different items on the checklist and mismatches between student and teacher choices were also recorded on an Excel spreadsheet. Qualitative data from both the survey and interviews were coded and analyzed thematically. All data were combined and closely analyzed to answer the research questions.

4. Results and Discussions

This section focuses on the findings of the research and discusses the implications of these findings.

4.1 The Accuracy of Students' Self-assessment

The average accuracy rate is 74.3%, meaning students assessed 23 out of the 31 items on the checklist correctly on average, with the highest accuracy rate of 100% and lowest 51.6%. Notably, a majority of students (36 in 57 students) achieved a more than 70% accuracy rate, although survey results showed that students were more confident about their SA, with 93.8% (61 out of 65 students) believing in their SA and chose different levels of "accurate". However, Participant 2 explicitly confessed "I am not very confident about my judgment" and Participant 6 in the interview admitted that she deliberately under-evaluated her performance when using the checklist ("I intentionally chose "No" for those I am not so sure about and I think I used a higher standard when I use the checklist"), showing a lack of confidence. Interestingly, students who reported "very accurate" SA in the survey actually mostly scored below average and those who reported "not very accurate" mostly scored above average (Table 1).

Table 1. Differences between perceived accuracy and actual accuracy

Very accurate		Not very accurate	
Participant	Score	Participant	Score
13	25	8	30
23	20	29	25
51	20	43	20
52	21	48	25

In terms of accuracy of different items, as is shown in Figure 1, students scored low on items related to language use (e.g. D2: Did you show a range of complex grammar structures (passive, clauses, etc.)?), referencing and citations (e.g. E1: Did your citations follow the XJTLU Harvard Referencing guide?) and academic style (e.g. F1: Did you mostly follow the academic style of writing?). Items with high accuracy are mostly those can make a decision at first glance (e.g. B2: Did you use sources?)



Figure 1. Accuracy rate of each item

Not only did the results above show common weaknesses among students (referencing mechanism and academic style rules), which teachers should address in teaching delivery, but also a gap between students' perceived SA and their actual accuracy and students showed different levels of confidence in their SA accuracy. This can be partially explained by students' lack of SA experience or training. In fact, although it is a common practice in the focal university that a checklist is offered to students for SA, it was reported in the survey that more than 20 students (33.7% of the survey participants) didn't use the checklist in Year 1 and 9.6% of participants seldom revised their work after using a checklist. To a certain extent, the numbers indicated a lack of awareness of using a checklist to compare their work with the desired product. Online teaching and relatively low participation because of the pandemic could be one of the reasons. A more possible reason behind this is maybe that Chinese students are more used to the judgment of authority such as a teacher rather than themselves and therefore lack relevant SA experience and training. This makes it more pressing and necessary that we increase the frequency of SA in class and SA training by modeling how to self-assess more effectively. According to Liu and Brantmeier (2019), SA experience and training partially influence SA accuracy and some evidence suggests practice or opportunities (Mahlberg, 2016) and training or instruction in SA through modeling (Kostons et al., 2012) can improve students' independent learning performance. Sung, Chang, Chiou, & Hou (2005) also found higher agreement between teacher and student assessments when students were trained in SA. Ross (2006) summarized different studies and concluded that students trained in SA in these studies consistently showed the reliability of SA across tasks and items and over short periods.

4.2 Factors that Influence Students' Accuracy of SA

Several major factors that tend to influence students' SA accuracy surfaced from the survey and interview results: item-related factors, lack of certain knowledge, timing, and attitude. First of all, it was reported by students that some items were very difficult to use. These items can be grouped into different types: those related to language use (sentence structure, academic vocabulary, etc.), those containing adjectives that were difficult to quantify (e.g. a range of, sufficient, etc.), those with terms and jargon (e.g. thesis statement, cohesive devices, etc.) and those containing words with a negative meaning (e.g. E4. Did you forget to include citations?). More specifically, survey results showed that the top three difficult items chosen by students were D2, D1, D3 (39.76%, 34.94% and 25.3% of participants respectively), all related to language use. In addition, E5, E3, E4 from Referencing and Citations category were rated difficult to assess among the most difficult half of items while items of other categories (except language use) seem to scatter. Meanwhile, two common difficulties were identified among interview participants: items with terms and items difficult to quantify. Regarding items with terms, Participant 6 explained that understanding of the terms such as "cohesive devices" would influence her judgment and "if I don't know the meaning, there will be a misunderstanding of the item". Participant 57 used the same example to explain the difficulty. As for items difficult to quantify, Participant 7 raised the question "how sufficient is sufficient?" and expressed her confusion. Secondly, results of low accuracy of items related to referencing and academic style showed a lack of awareness and knowledge in these areas. One respondent wrote in the survey that he/she lacked a proper understanding of references and citations. Thirdly, when to use the checklist was chosen to be an influencing factor. This was reflected in the survey results in particular that at least four participants confessed that after they have finished their writing for some time, they did not bother to read it again or they forgot what they wrote. Therefore, it is natural that their accuracy would be problematic when they cannot even remember the content. Finally, attitude towards SA also seemed to have influenced accuracy, which can partially be reflected by the "timing" factor and also a wrong judgment about those with a negative meaning (item E4, which was designed this way by the researchers deliberately) as it was not difficult to assess in theory.

Overall, there seems to be an overlap between different factors, not just the timing and attitude factors as explained above. Items about referencing and those with terms and jargon are similar in that both require students to notice beyond vocabulary and grammar and gain an awareness of subtle requirements of certain genre, which is a more profound requirement. However, we have to be realistic that some factors did influence students' SA and students were justifiable with a relatively low accuracy rate of those related to language use and those containing adjectives that were difficult to quantify. The reasons are: 1. language use related items are generally more difficult to evaluate as judgments of such items can be very subjective and it is more demanding on students' language proficiency. Even students with higher language proficiency were not promised with a high accuracy rate due to other factors, such as students' confidence or SA training for example. 2. As for those difficult to quantify, we as researchers had to discuss these items and standardize them before making a final decision.

The above-mentioned findings could have important implications for the improvement and design of checklists. Firstly, when designing the checklists, teachers will need to quantify and be as specific as reasonably possible. For example, phrases like 'a range of' could be replaced by a suggested number such as 'at least 5'. Secondly, we need to be careful with words containing negative meanings as they may also influence students' comprehension of relevant items and the accuracy of their SA. Thirdly, since many participants struggled with certain terms and jargon, it will be reasonable to attach a glossary of the terms used in the checklist as a supplementary document. This would help students to better understand the items on the checklist and at the same time, serve as a reviewing material to assist knowledge retention.

5. Conclusion

Through cross-examination of students' and teachers' checklists, a questionnaire, and interviews, the research intends to research the accuracy of Chinese students' SA in an EMI Chinese university and the factors that influence their accuracy. 65 Year 2 students participated in the study and their contribution has helped yield some meaningful findings. In terms of the first research question, results showed that the majority of students' SA were somewhat accurate although students showed different levels of confidence in their SA accuracy, which, along with students' lack of SA experience and training, has led to a discrepancy between their perceived and actual accuracy. Regarding the second research question, it was identified that factors such as checklist items, students' knowledge retention, timing, and attitude toward using the checklist can all influence their SA accuracy.

The following pedagogical suggestions were made: 1. Teachers should highlight certain components of academic writing in our teaching practices, especially those common weaknesses revealed in the checklist result. Teachers should provide further support or clarification to strengthen students' understanding of certain topics. This way, the use of a checklist serves its purpose of enhancing learning. 2. Teachers should provide SA training and modelling and create more SA opportunities. Considering the lack of relevant experience of Chinese students, it is also crucial that Chinese students are primed for the necessity and basics of SA, as according to Andrade (2011), understanding the value of SA is one of the key elements of SA. 3. Improving checklist design by designing more specific items and providing necessary support to assist the use of a checklist could also help increase SA accuracy.

Most of the findings such as the necessity of SA training and modelling are consistent with previous research results (Andrade, 2011; Bulter and Lee, 2010). However, the current study failed to establish a positive correlation between language proficiency and SA accuracy as mentioned above in the literature. It is a reasonable assumption that students with higher language proficiency should have a high SA accuracy as they have a better command of the language and should have a better understanding of items, especially those that appear to be difficult (e.g. items related to referencing and items with terms) as Dunning (2011) suggested that more competent learners are more able to accurately assess their work. Yet, with the influence of other factors identified in this study or maybe other sociocultural ones, this may be not always the case. Whether high language proficiency is promised with high SA accuracy could be a potential topic for future research.

References

- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48*(1), 12-19. <https://doi.org/10.1080/00405840802577544>
- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 13-18.
- Andrade, H. L. (2011). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade, & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). Nebraska: Routledge.
- Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education, 4*. <https://doi.org/10.3389/educ.2019.00087>
- Andrade, H. L., Mycek, K., & Du, Y. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy and Practice, 17*(2), 199-214. <https://doi.org/10.1080/09695941003696172>
- Ashton, K. (2014). Using self-assessment to compare learners' reading proficiency in a multilingual assessment framework. *System, 42*, 105-119. <https://doi.org/10.1016/j.system.2013.11.006>
- Bailey, K. M. (2003). *Learning about language assessment dilemmas, decisions and directions*. Beijing: Foreign Language Teaching and Research Press.
- Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me?: Individual self-assessment by skill and level of language instruction. *System, 40*(1), 144-160. <https://doi.org/10.1016/j.system.2012.01.003>
- Brown, G., & Harris, L. (2013). *Student self-assessment*. SAGE Publications, Inc., <https://doi.org/10.4135/9781452218649>
- Butler, Y. G., & Lee, J. (2010). The effect of self-assessments among young learners of English. *Language Testing, 27*(1), 5-31. <https://doi.org/10.1177/0265532209346370>
- Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice, 12*(1), 39-54. <https://doi.org/10.1080/0969594042000333904>
- Cheng, H., & Ding, Q. (2021). Examining the behavioural features of Chinese teachers and students in the learner-centered instruction. *European Journal of Psychology of Education, 36*, 169-186. <https://doi.org/10.1007/s10212-020-00469-2>
- Dann, R. (2002). *Promoting assessment as learning: Improving the learning process*. London: Routledge.
- Dunning, D. (2011). The Dunning-Kruger effect: On being ignorant of one's own ignorance. *Advances in experimental social psychology, 44*, 247-296. <https://doi.org/10.1016/B978-0-12-385522-0.00005-6>
- Eva, K. W., Cunnington, J. P. W., Reiter, H. I., Keane, D. R., & Norman, G. R. (2004). How Can I Know What I Don't Know? Poor Self-Assessment in a Well-Defined Domain. *Advances in Health Sciences Education: Theory and Practice, 9*(3), 211-224. <https://doi.org/10.1023/B:AHSE.0000038209.65714.d4>

- Kim, J. (2019). Effects of Rubric-Referenced Self-Assessment Training on Korean High School Students' English Writing. *English Teaching*, 74(3). <https://doi.org/10.15858/engtea.74.3.201909.79>
- Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, 22(2), 121-132. <https://doi.org/10.15858/engtea.74.3.201909.79>
- Lewis, J. (1990). Self-assessment in the classroom: A case study. In Geoff Brindley (ed.), *The second language curriculum in action* (pp. 187-213). Sydney: Macquarie University, National Centre for English Language Teaching and Research.
- Liu, H., Brantmeier, C. (2019). I know English”: Self-assessment of foreign language reading and writing abilities among young Chinese learners of English. *System*, 80, 60-72. <https://doi.org/10.1016/j.system.2018.10.013>
- Mahlberg, J. (2016). Formative self-assessment college classes improves self-regulation and retention in first/second year community college students. *Community College Journal of Research and Practice*, 39(8), 772-783. <https://doi.org/10.1080/10668926.2014.922134>
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2014). Rubrics vs. self-assessment scripts: effects on first year university students' self-regulation and performance / Rúbricas y guiones de autoevaluación: efectos sobre la autorregulación y el rendimiento de estudiantes universitarios de primer año. *Infancia y Aprendizaje*, 37(1), 149-183. <https://doi.org/10.1080/02103702.2014.881655>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74-98. <https://doi.org/10.1016/j.edurev.2017.08.004>
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research and Evaluation*, 11(10). <https://doi.org/10.7275/9wph-vv65>
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed). Upper Saddle River, NJ: Merrill/Prentice-Hall.
- Sung, Y.T., Chang, K. E., Chiou, S. K., & Hou, H. T. (2005). The design and application of a web-based self and peer assessment system. *Computers and Education*, 45(2), 187-202. <https://doi.org/10.1016/j.compedu.2004.07.002>
- Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*. London: Routledge.

Appendix A**WCW Business Evaluation Report Checklist**

Items	Yes/No
A. General	
A1. Did you use the proper report format (including sections with subheadings)?	
A2. Are most of the sources relevant?	
A3. Did you use different kinds of citations (integral, non-integral citations)?	
A4. Does each paragraph have a clear central topic?	
A5. Is information arranged logically? Is it easy to follow and understand?	
A6. Is it about one of the 3 industries?	
B. Introduction	
B1. Did you introduce the topic and the current situation?	
B2. Did you use sources?	
B3. Did you identify a market opportunity to exploit?	
B4. Did you include a thesis statement?	
B5. Did you state your business proposition in the thesis statement?	
B6. Did you specify the market analysis model in the thesis statement?	
C. Business proposition	
C1. Did you specify your product or service?	
C2. Did you name your product or service?	
C3. Did you specify the location of your business?	
C4. Did you specify your target customers?	
C5. Did you include future plans for your business?	
D. Language Use	
D1. Did you show a range of complex grammar structures (passive, clauses, etc.)?	
D2. Did you use a wide range of academic and business specific vocabulary?	
D3. Did you use proper collocations?	
D4. Did you use sufficient cohesive devices to better link ideas?	
E. Referencing and citations	
E1. Did your citations follow the XJTLU Harvard Referencing guide?	
E2. Did you use only the family name(s) of the author(s) in your citations?	
E3. Did you properly use punctuation? (such as () ‘ ’ , . or spaces)	
E4. Did you forget to include citations to acknowledge sources you have used?	
E5. Did you adequately paraphrase material from sources?	
F. Academic style	
F1. Did you mostly follow the academic style of writing?	
F2. Did you avoid 1 st /2 nd person pronouns such as ‘I’, ‘me’, and ‘you’?	
F3. Did you avoid colloquial language such as ‘kids’ or ‘boss’?	
F4. Did you avoid asking questions?	
F5. Did you avoid using contractions such as ‘don’t’, ‘can’t’, or ‘won’t’?	

Appendix B

Survey Questions

1. Have you used a checklist to evaluate your writing coursework in Year 1?

A. Yes. B. No

2. How useful do you think a checklist is?

A. very useful. B. somewhat useful. C. not very useful. D. not useful at all

3. Once you have ticked a "no" for an item, how often do you fix the problem in your writing later?

A. always. B. usually. C. seldom. D. never

If you choose option B, C, D in question 3, specify why

4. What items on the checklist do you think are difficult to use (e.g. A1, C3)? Can you list the items below:

5. How confident are you about the accuracy of your self-assessment using a checklist?

A. very confident B. somewhat confident

C. not very confident D. not confident at all

6. What factors do you think influence the accuracy of your self-assessment using a checklist, e.g. time spent on the checklist, understanding of the items, etc.

7. What did you do this time to be more accurate to use the checklist, e.g. highlighting all the citations or topic sentences, etc. (N.B. you can simply put down "nothing" if you did nothing extra when using the checklist to evaluate your work) ?

8. Do you have any suggestions on how we can help you better assess your work with a checklist?

Appendix C

Interview Questions

1. Do you think the checklist is useful? Why or why not?

2. Once you have ticked a "no" for an item, have you fixed the problem in your writing later? If no, why not?

3. On a scale of 0-10, how well do you think you understand all the items on the checklist?

4. What items do you think are difficult to use? How would you like them to be?

5. On a scale of 0-10, how confident are you about self-assessment? Why?

6. Do you have any suggestions on how we can help you better assess your work with a checklist?

7. When do you think would be a good time to use the checklist?

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).