JEELS
(Journal of English Education and Linguistics Studies)
P-ISSN: 2407-2575 E-ISSN: 2503-2194
https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/jeels

ASSESSING THE FEASIBILITY OF A WEB-BASED INTERACTIVE WRITING ASSESSMENT (WISSE): AN EVALUATION OF MEDIA AND LINGUISTIC ASPECTS

Hamamah [1]; Rafidah Sahar[2]; *Ive Emaliana[3]; Yulia Hapsari[4]; Putu Dian Danayanti Degeng[5]

[1]Department of Language and Literature, Brawijaya University, Indonesia; [2]Kuliyyah of Language Management, International Islamic University Malaysia, Malaysia; [3,5]Department of Language Education, Brawijaya University, Indonesia; [4]Department of English Language Education, Indonesia University of Education, Indonesia

hamamah@ub.ac.id; srafidah@iium.ac.id; ive@ub.ac.id; hapsari.yulia@upi.edu; dian_degeng@ub.ac.id*

(*) Corresponding Author

Abstract: This study aims to conduct a feasibility test on a Web-Based Interactive Writing Assessment (WISSE) that has been designed to assist students and teachers of Academic Writing classes by facilitating feedback provision online. WISSE was created to align with the requirements of the OBE curriculum. During this study, feedback about WISSE was gathered to enhance its feasibility and readiness. Trials focused on evaluating

[1]**Citation in APA style:**
Hamamah, Sahar, R., Emaliana, I., Hapsari, Y., & Degeng, P. D. D. (2023). Assessing the feasibility of a web-based interactive writing assessment (WISSE): An evaluation of media and linguistic aspects. *JEELS,* 10(1), 177-197.
DOI: 10.30762/jeels.v10i1.1093

Submission: March 2023, Revision: April 2023, Publication: May 2023

media aspects (usability, functionality, visual communication) and linguistic aspects (technical language, grammar, word choice). Quantitative and qualitative data were employed. The quantitative data were obtained through a material validation questionnaire and media validation using a rating scale on the quality of the media obtained from the objective assessment of media experts and linguists. The qualitative data were obtained from the feedback of media experts and linguists on the quality of the media provided in the comment column on the questionnaire. This data serves as a consideration in revising the product for the website's view and feasibility. the results of the expert test indicate that the web-based interactive learning assessment product, WISSE, is feasible for use, although revisions based on feedback from media and language experts are necessary before piloting it to a wider audience**.**

Keywords: *academic writing, feasibility test, interactive learning; web-based media, online writing assessment*

INTRODUCTION

Writing a scientific paper is a crucial skill for students and researchers in science-related fields who wish to share their discoveries and ideas with the academic community. However, students often struggle to produce high-quality scientific publications due to a lack of appropriate feedback from their instructors. Providing proper feedback is critical to the development of students' writing abilities. Graham (2019) emphasized the importance of providing timely and constructive feedback to improve students' writing abilities.

Kurnia (2022) noted that many students have difficulty understanding and incorporating feedback from their instructors, resulting in poor writing outcomes. Additionally, writing a scientific paper requires a solid understanding of scientific writing conventions, critical thinking, and research skills. Wulandari (2022) suggested that

effective feedback should help students recognize their strengths and weaknesses, provide specific improvement suggestions, and address issues such as grammar, clarity, and terminology.

Feedback plays a critical role in improving students' writing ability. Wisniewski et al. (2020) found that feedback on written tasks significantly influenced students' learning and engagement. However, one of the biggest obstacles in writing scientific papers is the lack of guidance and feedback from instructors on how to enhance writing outputs.

The conventional method of providing feedback, such as handwritten comments, can be time consuming and may not offer sufficient information. Moreover, grading students' writing skills through summative tests often fails to provide timely feedback. This lack of feedback can be significant challenge for students in improving their manuscript. Griffiths, et al. (2023) noted that feedback is an important factor in enhancing students' writing outcomes. Griffiths, et al. (2023) further suggested that feedback should be individualized, timely, and goal-oriented in order to assist students develop their writing abilities. Similarly, Sia & Cheung (2017) discovered that customized feedback is critical for students' learning and growth. Sia & Cheung (2017) noted further that students who receive individualized feedback can significantly enhance their writing outcomes.

To address these challenges, the development of interactive learning assessment that provides feedback to writers has been explored (Huang & Wilson, 2021). Computer-based feedback systems, offer more detailed and tailored feedback, which can aid students in improving their writing abilities (Kuklick, et al., 2023). These systems also allow for continuous evaluation and prompt feedback from teachers (Cavalcanti et al., 2021; Ramineni et.al, 2015). Cavalcanti et al. (2021) further explained that students who receive feedback through interactive learning assessment experience greater improvement in their writing abilities compared to those who do not. While this system improved scoring efficiency and provided quick feedback for most

types of questions, it had limitations in handling long-answer questions like essays. The available online writing applications often unable to accommodate corrections and unable to comprehend answers at the contextual level (Ni'am, Wibawa, & Endah, 2014; Ellis, 2009).

The demand for an effective feedback system in English academic writing in Indonesia is urgent, driven by various factors. The country seeks higher quality and increased quantity of scientific writing and publications (Minister of Research and Technology and Higher Education Regulation Number 50 of 2018). Indonesian writers face difficulties in writing scientific papers in English due to cultural differences in writing styles, as well as challenges with grammar, coherence, structure, and expressing ideas (Tahira & Haider, 2019; Wallace et al., 2004). Moreover, lecturers struggle to provide feedback and manage assignments due to large class sizes and limited time (Ariyanti & Fitriana, 2017; Hidayati, 2018; Juwariyah, 2015; Yustiana, 2015).

Although there are existing automated online applications that can assist writers in improving the quality of their writing, these tools have limitations such as limited accessibility, paid accounts, and a narrow focus on specific writing aspects. However, according to Bridgeman et al. (2012) and Monaghan & Bridgeman (2005), the scores generated by automated machines and human evaluators are generally comparable, exhibiting similar averages and standard deviations. These challenges and limitations highlight the need for further research and development of a comprehensive web-based writing assessment application (Hamamah et al., 2020).

In the study conducted by Yustiana (2015), a scoring system for Indonesian essays was developed by integrating the LMS Edmodo with techniques such as Latent Semantic Analysis and Euclidean Distance. The aim of the research described in this paper was to devise an alternative assessment system that caters to the particular requirements of writing in the Indonesian context, specifically considering the implementation of the Outcome Based Education

(OBE) curriculum. Building upon prior research (Hamamah et al., 2020), the WISSE website was developed as a platform to implement this assessment model. WISSE incorporates the framework of process writing and integrates the ability of multiple online writing applications to enhance the writing assessment process. WISSE also designed to utilized a combination of scores from the automated system and manual evaluations.

Evaluation of interactive learning assessments holds significant importance for several reasons. Firstly, it ensures the assessment's validity and reliability by examining its effectiveness in achieving learning objectives. Secondly, it aids in identifying areas for improvement and refining the assessment based on feedback and observed limitations. Thirdly, it evaluates the assessment's efficacy in promoting student learning, considering aspects such as performance, engagement, and motivation. Lastly, evaluation guarantees fairness by examining potential biases and accessibility challenges for students with diverse backgrounds and learning needs (Zhang, 2021; Liˇcen, et al., 2023).

Prior studies have underscored the significance of evaluation in interactive learning assessments (Elviana et al., 2020), highlighting its role in assessing efficacy, validity, and reliability. Therefore, the goal of this study is to evaluate the feasibility of utilizing WISSE, an interactive writing assessment tool to enhance students' scientific writing abilities. This study was conducted to continue the implementation and evaluation phases of the Analysis, Design, Development, Implementation, and Evaluation Model (Gagne et al., 2005), as well as to continue reporting the development of the web-based WISSE application prototype from previous research. The focus of this phase was on implementing WISSE to assess its feasibility for achieving product development goals and gathering feedback. The implementation trials were carried out to gather feedback, which would be utilized to make necessary enhancements and ensure the product is ready for widespread use. These trials focused on evaluating the media aspects, including usability, functionality, and visual

communication, as well as linguistic aspects such as the use of technical language, grammar, and appropriate word choice.

The feasibility study on WISSE was performed to address the following research inquiries: 1) What is the feasibility of WISSE as an interactive learning assessment accessible on the web? and 2) How reliable is WISSE as an interactive learning assessment accessible on the web?

METHOD

This study aimed to examine the feasibility of a Web-Based Interactive Learning Assessment product called WISSE applied in teaching English, especially in English-language academic writing. This study used both quantitative and qualitative data, where quantitative data was obtained through a questionnaire, and media validation, which was equipped with a rating scale for media quality, was assessed by media experts and linguists. The results of the assessment then went through an analysis process by averaging all the aspects of the assessment results to make sure that the media feasibility level was set up from the media and language aspects. As for the qualitative data, linguists and media experts provided responses and suggestions for the quality of the media provided in the comments column in the questionnaire. Inputs from these experts were used as considerations in revising the product.

In this case, the questionnaire used as a data collection instrument—where the researcher gave a set of questions or written statements to the respondents to answer (Sugiyono, 2008, p. 199)—had gone through the stages of validity testing before being used to test product usage. This questionnaire used a closed-ended questionnaire model, meaning that answer choices were provided for respondents to choose from. While the instrument validation test was intended to ensure that the instrument made was feasible to use and in accordance with the research objectives. This validity test was conducted on each

question item. The data obtained from validation by two media experts and two linguists were then analyzed using the average analysis technique, which finally resulted in a percentage of the feasibility level of the learning media being developed. Arikunto (2002, p. 216) explained that in order to determine the ranking of the final scores in each research questionnaire choice, the number of scores obtained must be divided by the number of respondents who answered the research questionnaire. The feasibility categories as presented by Arikunto and Jabar (2009, p. 35) are as follows:

Table 1. Feasibility level

Feasibility Assessment	Description
< 21 %	Very infeasible
21-40 %	Infeasible
41-60%	Moderate
61-80 %	Feasible
81-100%	Very feasible

By considering the expert judgment, the researcher subjectivity aspect related to research results, in this case the feasibility of a media, could be minimized. The results of expert judgment were processed through the stages of analysis with the Kappa coefficient approach to see the validity and reliability of the research results, considering that the research was qualitative. The validity of the success factors was conducted by asking the level of expert agreement based on a nominal scale questionnaire (2=agree, 1=disagree). Reliability tests were also conducted to compare and to see the consistency between the two experts (inter-rater agreement) in assigning a rating to the product, namely by measuring the Kappa index. Based on Fleiss (1981) through Viera et al (2005), the interpretation of Kappa values can be presented in Table 2.

Table 2. Kappa index

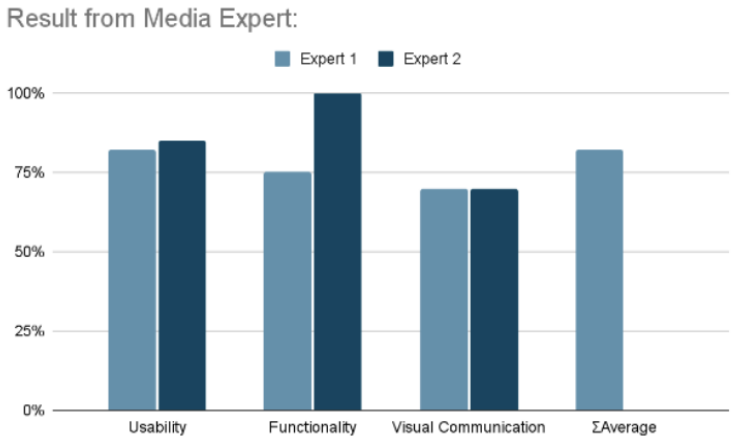
Kappa Coefficient	Agreement Level
< 0	None
0,01 – 0,20	Minimal
0,21- 0,40	Weak
0,41 – 0,60	Moderate
0,61 – 0,80	Strong
0,81 -0,99	almost perfect

By involving two different experts in testing WISSE, the researchers chose to tabulate the inter-rater agreement. Table 2 shows the magnitude of the Kappa coefficient and how to interpret the level of agreement between the two experts who assessed WISSE.

FINDINGS

Feasibility test result

The data obtained after filling out the questionnaire contained assessments and responses from two media experts and two linguists on the learning assessment media product in the form of the WISSE web. The data was divided into the feasibility of media aspects and linguistic aspects. The feasibility of the media aspects including usability, functionality, and visual communication with total 18 questions; while feasibility of the linguistic aspects including communication and language with total 6 questions; was validated by four experts from well-known universities, Universitas Brawijaya, Universitas Negeri Malang and Universitas Negeri Surabaya who were in charge of learning technology, held on 5 and 9 August 2021. The data results were shown in the following diagram illustration.

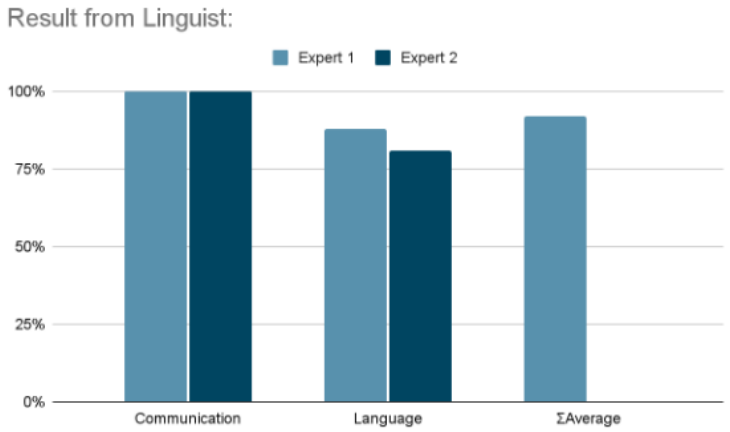


	Usability	Functionality	Visual Communication	ΣAverage
Expert 1	82%	95%	70%	82%
Expert 2	75%	100%	70%	
Average	79%	98%	70%	

Fig. 1. Result from media expert

As shown in Fig. 1, the percentage of feasibility assessment from the usability aspect of the two media experts obtained an average of 79.00%. While the functionality aspect obtained an average value of 98.00%. On the other hand, the visual communication aspect obtained an average of 70.00%. The total percentage of the three aspects obtained from the two media experts is 82.00%. According to the feasibility category as presented by Arikunto and Jabar (2009: 35), it can be concluded that WISSE's interactive media was described as very feasible to be used by a wide audience.

Meanwhile, the results of the validation test for linguistic aspects obtained from the responses and assessments of two linguists are shown in Fig. 2.



	Communication	Language	ΣAverage
Expert 1	100%	88%	92%
Expert 2	100%	81%	
Average	100%	84%	

Fig. 2. Result from linguists

Through Figure 2, the language feasibility of the two linguists obtained an average of 92.00%. Besides that, communication aspects obtained an average value of 100.00%. When averaged, the language aspect obtained a score of 84.00%. Meanwhile, 92.00% was the total percentage of aspects that influenced the language feasibility test. Based on the feasibility category (Arikunto, 2009: 35), the linguistic aspects in the WISSE assessment media product met the established standards. After receiving the results of the assessment and feedback from media and language experts, the next step was to analyze and revise the product according to the comments that had been given.

Inter rater reliability result

To see the consistency between two experts (raters) in assessing WISSE products, SPSS Statistical Software Version 22 was used to determine the magnitude of the Kappa coefficient to calculate inter-

rater reliability. Table 1 shows the consistency of judgments between experts. Of the five aspects assessed, all (72.2%) have consistent values.

Tabel 1. Rater_1 * Rater_2 cross tabulation

		.00	1.00	Total
Rater_1	.00	.22.2%	11.1%	33.3%
	1.00	5.6%	61.1%	66.7%
Total		27.8%	72.2%	100.0%

Table 2. Kappa coefficient

	Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Kappa Agreement	.609	.202	2.605	.009
N of Valid Cases	17			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

From table 2, the reliability coefficient between raters (Kappa) was 0.609. Based on table 2, the Kappa coefficient is interpreted in the strong agreement category because the Kappa coefficient value is > 0.60. The five aspects contained in the questionnaire have been tested for validity and reliability. Therefore, the instrument meets the reliability requirements.

DISCUSSION

The success in the learning process could not be separated from the existence and novelty of facilities provided by a university, especially in each faculty according to the needs of both students and lecturers who continued to grow (Yulando et al., 2019). This argument was linear with several aspects influencing the success of the learning process, such as rapid technological developments, demands to complete assignments on time, and demands to be able to provide flexible and two-way educational guidance. The role of technology was

influential to be applied in an all-high-tech environment, especially when the learning processes had switched to online learning, where students and lecturers did not need to be in the same place in conducting the learning process. Therefore, this study aimed to bridge the problems faced by students and lecturers by utilizing technological sophistication, especially when they were faced with challenges of writing scientific articles in English. These challenges could be in the form of a lack of quantity and quality of feedback given by lecturers to improve manuscripts before publication.

Koppi et al (1997) explained that there were two main aspects that technology could offer to increase the success of the learning process, especially in providing feedback. First, the technology used must mimic or represent real-world interactions as much as possible where experiential learning (experience-based learning) was realized, as well as in collaborative learning that used site-based communication facilities. This was the functional role of the WISSE web that was developed. To make this happen, WISSE had passed through several evaluation phases by experienced experts in their fields, such as media experts and linguists. Two media experts had the opportunity to examine the interface and performance of the WISSE web prototype, while two linguists observed the proper use of technical language, grammar and word choice for a learning site. Apart from filling out the questionnaire as described in the previous chapter, experts also provided opinions and suggestions to developers and users in relation to the WISSE web.

From the results of the questionnaire that the media experts filled out after evaluating the WISSE site, it was concluded that the three aspects or the main assessment aspects consisting of usability, functionality, and visual communication were all considered very feasible to operate as learning media. This means that both media experts agreed that the WISSE site was an informative and interactive site and could be accessed easily, where all the menus—such as the "Home", "News", "Tools" menus, and so on—function properly. They also agreed that a site design including color choices, layout, and

typography was unambiguous and easy to navigate. The feedback provided by media experts 1 revolved around the size of the WISSE site name's font size being small and less prominent, as well as the lack of a logo as a site identity which was considered significant enough for future improvements.

In this case, the importance of enlarging the font size of WISSE's site name was considered to be followed up immediately because most visitors focused their attention for the first time on the logo, headers, and site title listed on the site's main page (Yesilada et al., 2008). This was in line with the research findings of Grobelny and Michalski (2015) where large and tight text sizes attracted more attention than small and spaced texts. It was also possible that media experts 1 suggested that the name WISSE be enlarged so that this site gave a certain impression and be easily remembered by users and visitors, even though human memory was not completely affected by font size alone (Mueller et al., 2014).

Media expert 1 also regretted that this site had not yet had a unique and identical logo as part of the identity of the WISSE site. Adir et al (2012) emphasized that the logo was an important element consisting of unique and special details as a differentiating aspect that distinguishes one site from another. The logo, which was also the visual aspect that first came to the attention of users or visitors, could influence their positive responses and attitudes towards the performance of the interface and the overall design of a site, especially when the logo design was more attractive and complex (Machado et al., 2015; van Grinsven & Das, 2016). This was why the existence of a logo for a new site like WISSE had such a big impact on attracting visitors and giving regular users comfort when using this site for their various academic writing needs.

On the other hand, linguists concluded that the WISSE site was very feasible for mass operation. This could be seen from the easy-to-understand language and communication aspects, where the two experts agreed that spelling, grammar, word choice, and the use of punctuation marked on the WISSE website followed good and correct

language rules. Linguists, especially linguists 2, also paid attention to the function of several language features such as grammar and plagiarism checkers that had been embedded in the WISSE web containing more words than other similar features. Some of the comments and suggestions from the two experts revolved around choosing inappropriate diction, the function of typeface emphasis—such as italics, bold, and underline—which was not appropriate, and the effectiveness of sentences to avoid repetition.

Intrinsically, the convenience of navigating a site depended on four aspects of user and visitor behavior when browsing a site, namely distribution of attention, confidence in choice of link, efficiency, and effectiveness (Katsanos et al., 2010). Knowing that navigating or exploring a site was driven by certain goals, it was important for a site to direct its users and visitors to the specific page they really wanted to go to by choosing the right choice of diction that effectively and efficiently represented the content of that page (Gillis, 2017 ; Katsanos et al., 2010). User experience, in this case, was essential to make sure they did not get "lost" when browsing the site caused by inaccurate word choices by the site development team. This was also related to sentence composition that was less effective, made a lot of unnecessary repetition and caused confusion. Especially when the target users of the WISSE site were students who might still be unfamiliar with education-based sites with all their technical terminology. When observing the academic library website at a university in Canada, Gillis (2017, p. 20) emphasized that students tended to quickly misinterpret the function of a feature in their library site. The choice of words interpreted differently by each person was actually the impact of the users' own browsing behavior in cyberspace, where one word had different meanings from one site to another. This was the reason why linguists emphasized that word choice—such as replacing "newest first" with "latest" with the assumption that the last word was widely used by other sites students were familiar with—was very important for immediate follow-up on future improvements.

As for typeface emphasis, there was a possibility that linguists wanted the WISSE site development team to be more careful in using the right typeface by paying attention to its function and its effect on sentences stressed—whether it emphasized italics, bold, or underlined. Brumberger (2003) summarized the prevailing opinions that typography, or typeface, had its own personality, and many practitioners assigned a certain persona to certain types of typefaces. This is done because readers behaved differently when faced with italicized or bold text. In a previous study that observed typographical differences in text, Dyson and Beier (2016) emphasized that changing typefaces from bold to neutral and bold to italic had different effects on word and sentence identification. In this case, linguist 1 advised the WISSE web development team to pay attention to the use of italics in the sentence "The instructor is authorized to open…". This could imply that the linguist 1 wanted the site's development team to weigh the use of italics that was generally known to reduce reading speed and word recognition rates, although it was not certain whether this was due to the less legible italics features, or because indeed the reader was not familiar with the emphasis on italics in the text (Bigelow, 2019).

CONCLUSION

In learning foreign languages, interactive learning assessment that utilized the convenience of technology and were interactive were currently one of the needs of lecturers to overcome various problems in the teaching and learning process and to motivate students to use the technology carefully and precisely. Therefore, interactive and educative multimedia products continued to be developed to bridge the barriers experienced by both students and lecturers. Therefore, this study aimed to test the feasibility of a web-based interactive learning assessment product called WISSE designed not only to help teaching English—especially in terms of academic writing—but it was hoped that this product could also be used for other academics who had the same problems when required to write academic in English. This feasibility test focused on assessing media aspects—such as usability,

functionality, and visual communication—and linguistic aspects—such as the use of technical language, grammar, and the right choice of words—a WISSE product involving media and language experts from well-known universities. From the results of a questionnaire that had been tested for validity and reliability, media and language experts agreed that this WISSE interactive assessment product was suitable to use and test on academics, such as lecturers and students. These experts also provided some specific suggestions and inputs related to media and linguistic aspects that needed to be reconsidered for the benefit of future product development. After revising the interactive assessment product according to the results of the assessment and feedback from language and media experts, it was hoped that the WISSE application could soon be mass-tested to help academics who had problems teaching English and academic writing in foreign languages.

ACKNOWLEDGMENTS

This work was supported by Research and Community Service Grant of Universitas Brawijaya, Indonesia.

REFERENCES

Adir, G., Adir, V., & Pascu, N. E. (2012). Logo design and the corporate identity. *Procedia - Social and Behavioral Sciences*, *51*, 650–654. https://doi.org/10.1016/j.sbspro.2012.08.218

Arikunto, S. (2002). *Metodologi penelitian suatu pendekatan proposal*. Rineka Cipta.

Arikunto, S., & Jabar, C. S. A. (2009). *Evaluasi program pendidikan*. Bumi Aksara.

Ariyanti, A., & Fitriana, R. (2017). EFL students' difficulties and needs in essay writing. *Proceedings of the International Conference on Teacher Training and Education 2017 (ICTTE 2017)*. https://doi.org/10.2991/ictte-17.2017.4

Prior, M. T. (2019). Elephants in the room: An" affective turn," or just feeling our way?. *The Modern Language Journal, 103* (2), 516-527.

https://doi.org/10.1111/modl.12569

Bridgeman, B., Trapani, C., & Attali, Y. (2012). Comparison of human and machine scoring of essays: Differences by gender, ethnicity, and country. *Applied Measurement in Education*, *25* (1), 27–40. https://doi.org/10.1080/08957347.2012.635502

Brumberger, E. (2003). The rhetoric of typography: The persona of typeface and text. *Technical Communication (Washington)*, *50* (2), 206–223.

Cavalcanti, A.P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y., Gasevic, D., Mello, R.F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence 2.* https://doi.org/10.1016/j.caeai.2021.100027

Dyson, M. C., & Beier, S. (2016). Investigating typographic differentiation: Italics are more subtle than bold for emphasis. *Information Design Journal*, *22* (1), 3–18. https://doi.org/10.1075/idj.22.1.02dys

Ellis, R. (2009). Corrective feedback and teacher development. *L2 Journal, 1* (1). doi:10.5070/l2.v1i1.905410.5070/L2.V1I1.9054 downloaded on 2020-02-21

Elviana, R., Inderawati, R., & Mirizon, S. (2020). Developing interactive multimedia for teaching descriptive texts-based on Palembang local culture. *Indonesian EFL Journal, 6* (1), 19-26. doi: 10.25134/ieflj.v6i1.2634.

Fleiss, J. (1981). *Statistical methods for rates and proportions* (1st ed.). John WIley & Sons.

Gagne,R.M., Wager, W.W., Golas, K., Keller, J.M. (2005) *Principles of instructional design (5th ed.)*, Thomson: Wadsworth.

Gillis, R. (2017). "Watch your language!": Word choice in library website usability. *Partnership: The Canadian Journal of Library and Information Practice and Research*, *12*(1). https://doi.org/10.21083/partnership.v12i1.3918

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education 43* (1): 277-303.

Griffiths, C.M., Murdock-Perriera, L., Eberhardt, J.L. (2023). "Can you tell me more about this?": Agentic written feedback, teacher expectations, and student learning. *Contemporary Educational Psychology 73*. https://doi.org/10.1016/j.cedpsych.2022.102145

Grobelny, J., & Michalski, R. (2015). The role of background color, interletter spacing, and font size on preferences in the digital presentation of a product. *Computers in Human Behavior*, *43*, 85–100. https://doi.org/10.1016/j.chb.2014.10.036

Hamamah, H., Hapsari, Y., Emaliana, I., & Degeng, P. D. D. (2020). Integrated academic writing assessment model to support the implementation of OBE curriculum. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, *5*(7), 1040. https://doi.org/10.17977/jptpp.v5i7.13827

Hidayati, K. H. (2018). Teaching writing to EFL learners: an investigation of challenges confronted by Indonesian teachers. *Langkawi: Journal of The Association for Arabic and English*, *4*(1), 21. https://doi.org/10.31332/lkw.v4i1.772

Huang, Y. & Wilson, J. (2021). Using automated feedback to develop writing proficiency. *Computers and Composition 62*. https://doi.org/10.1016/j.compcom.2021.102675

Juwariyah, S. (2015). Penilaian dalam pembelajaran menulis bahasa Indonesia. *Magistra*, *93*, 73–84.

Katsanos, C., Tselios, N., & Avouris, N. (2010). A survey of tools supporting design and evaluation of websites based on models of human information interaction. *International Journal on Artificial Intelligence Tools*, *19*(06), 755–781. https://doi.org/10.1142/S0218213010000418

Koppi, A. J., Lublin, J. R., & Chaloupka, M. J. (1997). Effective teaching and learning in a high-tech environment. *Innovations in Education & Training International*, *34*(4), 245–251. https://doi.org/10.1080/1355800970340402

Kuklick, L., Greiff, S., & Lindner, M. A. (2023). Computer-based performance feedback: Effects of error message complexity on cognitive, metacognitive, and motivational outcomes. *Computers*

& Education, 200, 104785.

Kurnia, A. (2022). Efl Students' Problems In Dealing With Teacher Written Feedback. *Jurnal Pendidikan Indonesia*, *3*(05), 485–492. https://doi.org/10.59141/japendi.v3i05.1143

Liˇcen, S.; Cassar, M.; Filomeno, L.; Yeratziotis, A.; Prosen, M. (2023) Development and Validation of an Evaluation Toolkit to Appraise eLearning Courses in Higher Education: A Pilot Study. *Sustainability 15, 6361*. https:// doi.org/10.3390/su15086361.

Machado, J. C., de Carvalho, L. V., Torres, A., & Costa, P. (2015). Brand logo design: Examining consumer response to naturalness. *Journal of Product & Brand Management*, *24*(1), 78–87. https://doi.org/10.1108/JPBM-05-2014-0609

Monaghan, W., & Bridgeman, B. (2005). E-rater as a quality control on human scores. *R&D Connections*, *2nd*.

Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's beliefs about memory? *Journal of Memory and Language*, *70*, 1–12. https://doi.org/10.1016/j.jml.2013.09.007

Ni'am, S., Wibawa, H.A., & Endah, S. N. (2014). Pengembangan aplikasi Learning Management System (LMS) pada Sekolah Menengah Pertama Islam Terpadu (SMP IT) Harapan Bunda Semarang, Journal of Informatics and Technology, vol. 2, no. 1, pp. 11-32, Sep. 2014. [Online]. Retrieved from: https://ejournal3.undip.ac.id/index.php/joint/article/view/6266/0

Ramineni, C., Trapani, C.S., Williamson, D.M. (2015). *Evaluation of e-rater for the Praxis I Writing Test.* doi:10.1002/ets2.12047

Sia, D. & Cheung, Y.L. (2017). Written corrective feedback in writing instruction: A qualitative synthesis of recent research. Issues in Language Studies 6(1): 61-80.

Sugiyono. (2008). *Metode penelitian kuantitatif kualitatif dan R&D*. Alfabeta.

Tahira, M., & Haider, G. (2019). The role of critical thinking in academic

writing: An investigation of efl students' perceptions and writing experiences. *International Online Journal of Primary Education*, *8*(1), 1–30.

van Grinsven, B., & Das, E. (2016). Logo design in marketing communications: Brand logo complexity moderates exposure effects on brand recognition and brand attitude. *Journal of Marketing Communications*, *22*(3), 256–270. https://doi.org/10.1080/13527266.2013.866593

Viera, Anthony,J., and Joanne M. Garrett. (2005). Understanding interobserver agreement of the kappa statistic. Fam Med 37.p: 360-363.

Wallace, T., Stariba, W. E., & Walberg, H. J. (2004). Teaching speaking, listening and writing. *International Academy of Education*, *May*, 13.

Wisniewski, B., Zierer, K., Hattie, J. (2020). The power of feedback revisited: a meta-analysis of educational feedback research. *Frontiers in Psychology 10.* https://doi.org/10.3389/fpsyg.2019.03087

Wulandari, Y. (2022). Effective feedback to improve students' writing skills. *English Education, Linguistics, and Literature Journal 1*(1): 10-17

Yesilada, Y., Jay, C., Stevens, R., & Harper, S. (2008). Validating the use and role of visual elements of web pages in navigation with an eye-tracking study. *Proceedings of the 17th International Conference on World Wide Web - WWW '08*, 11. https://doi.org/10.1145/1367497.1367500

Yulando, S., Sutopo, S., & Franklin Chi, T. (2019). Electronic module design and development: an interactive learning. *American Journal of Educational Research*, *7*(10), 694–698. https://doi.org/10.12691/education-7-10-4

Yustiana, D. (2015). Penilaian otomatis terhadap jawaban esai pada soal berbahasa Indonesia menggunakan latent semantic analysis. *Seminar Nasional "Inovasi Dalam Desain Dan Teknologi" - IDeaTech 2015*, 123–130.

Zhang, S. (2021). Review of automated writing evaluation systems.

Journal of China Computer-Assisted Language Learning, *1*(1), 170-176. https://doi.org/10.1515/jccall-2021-2007