

How mixed reality shifts visual attention and success in experimental problem solving

Dörte Sonntag^{*} and Oliver Bodensiek[†]*Institute for Educational Research in Natural Sciences, Technische Universität Braunschweig,
Bienroder Weg 82, 38106 Braunschweig, Germany*

(Received 1 October 2021; accepted 8 June 2022; published 21 July 2022)

While commercially available mixed-reality (MR) head-mounted devices are also increasingly used in education the impact of MR learning environments is mostly being evaluated with respect to learning outcomes and learning gains alongside a number of affective variables. Here we aim at a deeper understanding of the influence of MR on experimental problem-solving processes by means of eye-tracking analyses as an addition to self-reported and questionnaire-based test instruments. On the basis of a didactically and empirically sound designed MR learning environment for experimental problem solving with an electrical circuit we investigate how novices and experts differ in their success in problem solving and their visual attention for both a MR learning environment and a non-MR condition. Our results indicate that MR not only has a positive effect for both groups on success in problem solving, but that it is also accompanied by significant shifts of visual attention on the different parts of the experiment.

DOI: [10.1103/PhysRevPhysEducRes.18.023101](https://doi.org/10.1103/PhysRevPhysEducRes.18.023101)

I. INTRODUCTION

Mixed reality (MR) and corresponding MR learning environments are of great interest for science education research as they offer entirely new opportunities for and approaches to teaching and learning [1,2]. Most of the research in this field, however, is still limited to tablet-based augmented reality (AR), even though MR via head-mounted devices (HMD) offer advanced possibilities to support learners in lab-based learning activities: Wearing a HMD allows for hands-free experimentation, the virtual representations or simulations can be displayed and perceived three dimensionally and superimposed on the physical environment more naturally, and an intuitive gesture-based interaction with the virtual elements is possible. MR-HMD are thus capable to combine learning in the laboratory with the advantages of interactive computer simulations such as PheT [3,4]. The few studies that explicitly use MR-HMD indicate a positive impact with respect to conceptual knowledge [5–7] and also improvements of performance in mechanical assembly [8] and other procedural tasks [9].

Recent commercially available MR-HMD such as the Microsoft HoloLens 2, which is used throughout this study, are equipped with integrated eye-tracking functionalities

and thus offer an easy-to-use opportunity to evaluate MR applications [10] with respect to visual attention and eye-tracking metrics that are accessible within the hardware limitations [11]. Eye-tracking data are widely used in the human-computer interaction context, e.g., for usability evaluation studies or to guide system interaction [12]. It is nowadays also widespread in physics education research with respect to, e.g., problem solving [13] or utilization of multiple representations in reasoning [14].

The approaches of novices and experts to experimental problem solving often differ fundamentally. This is partly related to the different amount of procedural knowledge in scientific work, but to a large extent it is also due to prior subject specific knowledge and the use of visual and mental representations [15–17]. The whole experiment-related problem-solving process can be summarized in four main phases: Finding suitable representations, selecting a solution method, implementing the method, and evaluating it [18,19]. Experts often take more time to form representations in the first phase, while novices complete this phase quickly by using rather naive representations without any relation to domain-specific content [20]. Hence, this very first phase should also be specifically addressed in MR by offering tools for building problem-related representations.

The benefits of MR in problem-solving tasks have as yet not been studied extensively. Although there have been a number of studies showing a positive effect of MR on spatial problem solving [21], this does not directly relate to the type of problem solving under consideration here. In the area of (tablet-based) AR, studies show a positive impact of technology on students' confidence in their problem-solving skills [22], which may indicate how MR might help in this direction as well.

^{*}Corresponding author.

doerte.sonntag@tu-braunschweig.de

[†]o.bodensiek@tu-braunschweig.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

II. STUDY DESIGN AND METHODS

Our study aims to test the hypothesis that MR supports subjects with low prior knowledge, which we refer to as novices in the following, to achieve higher success in experimental problem solving. Participants with high prior knowledge, in turn, are referred to as experts. In order to elicit prior knowledge that is relevant for the experiment in this study and to divide participants in novices and experts, a literature-based and quality-criteria-checked multiple-choice questionnaire was developed, which participants completed first.

For the experiment a rather unknown electric circuit was chosen so that neither novices nor experts normally would be familiar with the setup and immediately know how to approach it. The electrical circuit consists of a couple of different elements as shown in Fig. 1(a). Participants are given two tasks that are based on a rather general two-phase structure of the experimental problem-solving process. In the first task, participants are asked to get an overview of the circuit and understand its functions, which is mainly based on preknowledge including, e.g., facts and concepts. To do so, both switches can be used by the participants to experimentally gain understanding and match their mental models with experimental observations. In the second task the focus is on the execution of the problem-solving activity. The aim is to maximize the time between the lower lamp going out and the upper lamp lighting up. For this purpose, participants are given additional building blocks such as coils, resistors, and capacitors with different characteristic physical parameters.

Subjects were randomly assigned either to a mixed reality learning environment (MRLE) or to the standard condition without using MR. In the non-MR condition a Pupil-Labs head-mounted eye-tracking device (ETD) is worn throughout the experiment, while in the MR setting a HMD is worn during experimentation [Fig. 1(b)] such that additional virtual information can be superimposed on the circuit.

With respect to the experimental setup, the two conditions differ only in that additional multimeters to measure voltage or current are given for optional use in the non-MR

condition, while a similar function is implemented directly in the MRLE: Real-time voltages are sent to the MR-HMD during the experiment for the visualizations to adjust accordingly. The display of the according values could be optionally switched on in the MRLE. In both settings the use of virtual or physical measuring apparatus is optional and hence the intensity of usage varies strongly among the participants. The voltage drops at the lamps were measured most frequently.

The design of the MRLE was derived from an eye-tracking based design process as proposed in Ref. [23] and features visualizations of the electric current on the wires [Fig. 1(c)], a charge level indicator of the capacitors, current-time and voltage-time curves to trace exactly the operation of the capacitors and the coil and further visualizations such as formulas, potential curves, magnetic fields, or a timer function. All visualizations can be selected or deselected in the MRLE individually by the users as needed.

In order to test if the MRLE supports in problem solving, the according success is quantified along the following grading scheme: A maximum of one point is awarded for each relevant module (coil, resistor, and capacitor) that has been replaced in the best possible manner. If replacement only improved the result but did not reach the optimum, only half of a point is awarded. The capacitors in the middle branch can be considered to illustrate this case: If the capacity has been reduced compared to the initial state, such that the lower lamp turns out faster, but has not been selected in the best possible way, 0.5 points are awarded for this module.

During the experiment, eye-tracking data of the participants are recorded either by the ETD or directly by the MR-HMD. Both eye trackers are calibrated prior to use for each participant. The ETD provides a somewhat better spatial resolution of 0.6° spatial angle compared to 1.5° of the HMD. With respect to the distance of about 1 m from the eyes to the experimental setup, however, both result in an accuracy of the same order of magnitude of a few centimeters, which lies well below the size of the areas of interest (AOIs). The latter are defined for statistical analysis of eye-tracking data and illustrated in Fig. 1(a).

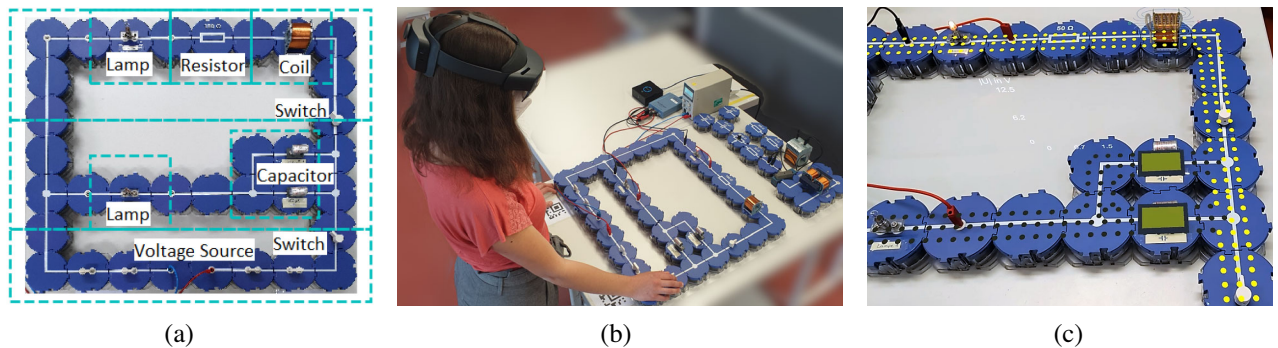


FIG. 1. Experimental setup. (a) Experiment setup: AOIs are indicated in light blue, (b) MR-setup: participants wear MR-HMD throughout the experiment, and (c) View through the MR-HMD on the experiment.

TABLE I. Achievements in the prior-knowledge questionnaire of the four independent groups.

		<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>T</i>	<i>df</i>	<i>p_{corr}</i>
Non-MR	Novices	9	12.7%	5.1%	5.71	17	0.005**
	Experts	10	40.4%	13.7%			
MR	Novices	13	11.6%	5.6%	6.23	15	0.005**
	Experts	13	40.4%	15.6%			

Significance level: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

III. RESULTS

A. Sample

The sample consists of $N = 45$ university students from science, technology, engineering, and mathematics programs, both undergraduate and graduate level, equally distributed to both conditions. According to Sec. II the participants are assigned to four independent groups. The questionnaire results with respect to prior knowledge are summarized in Table I for each of these groups. Novices and experts differ significantly [24] in both conditions while there is no significant difference for experts or for novices between the MR and non-MR conditions. Therefore it is assumed that both experts and novices are from the same statistical population.

B. Problem-solving success

Although neither both novice groups, nor both expert groups differ significantly in their prior knowledge, it can be seen that novices as well as experts improved their problem-solving success significantly in the MR condition: Novices improved by 39.7% and experts still by 23.6% with huge corresponding effect sizes according to Cohen (cf. Table II) [25], where d indicates a measure by how many standard deviations the means differ, i.e., here a substantial difference of more than 2 standard deviations can be seen for novices and even more than 1 standard deviation for experts. This is a clear evidence for improvement of problem-solving success by using the MRLE.

A two-way ANOVA shows no interaction-effect between the two independent categorical variables learning environment (MR, non-MR) and preknowledge (experts, novices) on the continuous variable problem-solving success

TABLE II. Problem-solving success of the four independent groups.

		<i>Mean</i>	<i>SD</i>	<i>T</i>	<i>df</i>	<i>p_{corr}</i>	<i>d</i>
Novices	Non-MR	33.3%	25.0%	4.29	12	0.005**	2.05
	MR	73.1%	14.5%				
Experts	Non-MR	53.3%	23.3%	2.62	13	0.04*	1.33
	MR	76.9%	18.7%				

Significance level: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

[$F(1, 41) = 1.8$, $p = 0.20$, partial $\eta^2 = 0.04$]. The homogeneity of variance is given by Levene [26]. However, there is a strong significant effect for the environment on problem-solving success ($p_{\text{corr}} < 0.005$).

C. Visual attention

In the following eye-tracking-based analysis we restrict ourselves to the first experimental phase, since in this phase the experiment is to be understood by the participants and it thus forms the most important basis for further problem-solving processes.

A qualitative analysis of differences between novices and experts both in the MR and non-MR condition are compared using heatmaps showing the spatial distribution of sample-averaged relative gaze rates. We will refer to the latter as *gaze rates* in the following. Based on the integrated gaze rates in each AOI, further quantitative statistical analysis is performed to compare the novices' approaches with experts' approaches within each of both MR and non-MR conditions, and novices' and experts' approaches between the MR and the non-MR condition, respectively.

The heatmaps of the non-MR setting are shown for comparison in Figs. 2(a)–2(b). The most distinctive gaze rates of both novices and experts are found on some of the relevant circuit components, such as capacitors and lamps. Only little attention is paid to the coil and the resistor in both groups. Novices are strongly focused on the elements of the middle branch of the circuit and little attention is paid to the upper branch, while experts' visual attention is more balanced. This could be because the middle lamp turns off for no visual reason, since one cannot see the charging of the capacitors. Overall, it may seem that novices focus more on relevant elements of the circuit, while experts focus more on the entire circuit. This is mainly due to the higher transition rate of experts compared to novices, as they tend to switch between elements more frequently, while novices tend to focus on one element for longer.

Considering the heatmaps for the MR setting in Figs. 2(c)–2(d), it is important to note that the circuit setup is slightly different, as the coil was moved from left to right and the resistor from the right to the center on the top branch, which was necessary due to the coupling of the measurement device to the circuit in the MRLE. However, this is not a detriment to the analysis of the AOIs.

The visual attention in the MRLE is more focused on the circuit elements and not as broadly scattered as in the non-MR condition. This effect may be partly due to the different (practical) resolutions of the eye-tracking devices, but also the MRLE may likely direct the participants' gaze more onto the circuit elements. By further comparing non-MR and MR conditions, increased gaze rates along the wires especially among the novices in the MRLE can be noticed. In the non-MR condition we could mainly recognize this phenomenon in the experts' gaze patterns, specifically in the analysis of their gaze paths (not shown

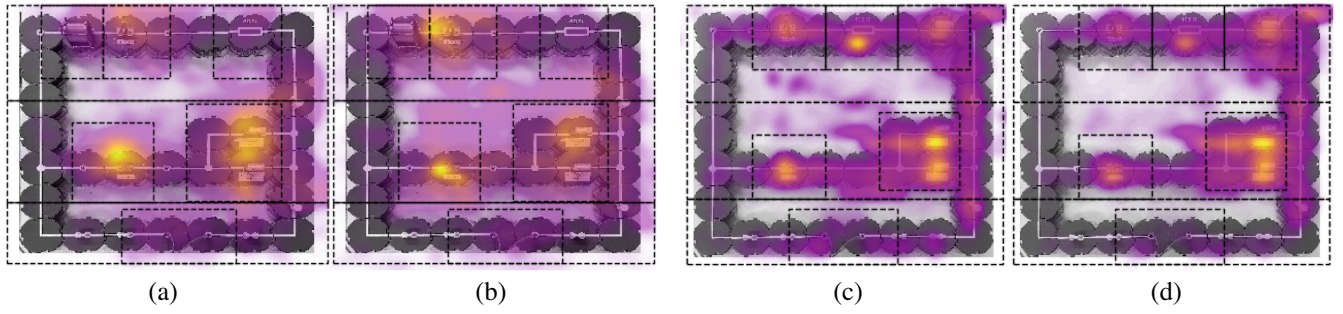


FIG. 2. Group-averaged relative gaze rate distributions for all four groups. Each participant is weighted equally. The important AOIs are indicated in gray. (a) Novices, non-MR, (b) Experts, non-MR, (c) Novices, MRLE, and (d) Experts, MRLE.

here). Furthermore it can be noticed that novices in the MRLE pay more attention to the upper circuit than experts. Interestingly, this has reversed compared to non-MR condition.

In the following, groupwise cumulative visual attention based on the previously defined AOIs is analyzed using a MANOVA [27]. The homogeneity of the error variances is met for all considered variables as Levene's test shows no significance. Additionally the homogeneity of covariances is given, as assessed by Box's test ($p > 0.001$). The further requirements are also met, but in some cases, such as the normal distribution or linearity of some groups in some variables only with restrictions, but still good enough to proceed [27]. The MANOVA shows statistically significant differences with a large effect (cf. $\eta^2 > 0.14$) between the groups for the combined dependent variables with $F(24, 99) = 7.744$, $p < 0.001$, partial $\eta^2 = 0.640$ and Wilk's $\Lambda = 0.047$. The strong significance accompanied by a low Wilk's Λ allows us to reject the null hypothesis.

TABLE III. *Post hoc* analysis from the MANOVA.

	F	df_1	df_2	p	η^2
Resistor	2.669	3	41	0.06	0.163
Coil	3.157	3	41	0.035*	0.188
Capacitors	1.003	3	41	0.401	0.068
Lower lamp	6.749	3	41	<0.001**	0.331
Upper lamp	7.417	3	41	<0.001**	0.352
Middle circuit	9.105	3	41	<0.001**	0.400
Upper circuit	6.201	3	41	0.001**	0.312

Significance level: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

TABLE IV. Significant between-group comparisons between non-MR and MR conditions.

	Group	$\Delta\mu$	SE	p
Resistor	Novices	0.0169	0.006	0.048*
Lower lamp	Novices	-0.07	0.023	0.022*
Lower circuit	Novices	-0.176	0.042	0.006**
Upper lamp	Experts	-0.073	0.018	0.008**

Significance level: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

To identify the associated variables and groups to the significant differences, a *post hoc* analysis was performed. We found significant differences in the gaze rate distributions on the resistor, the coil and both lamps with a large effect (cf. Table III). For the groups of novices we could identify the gaze rates on the resistor and the lower lamp to be significantly different between non-MR and MR conditions (cf. Table IV). While the gaze rate on the lower lamp decreases for the novices it decreases for the experts on the upper lamp. This is consistent with our previous findings from the heatmap analysis.

IV. DISCUSSION

The problem-solving success of novices improved significantly with a huge effect in the MRLE as compared to the non-MR condition, supporting our hypothesis from Sec. II. This might stem from the fact that they get support from tools and visual representations in the MRLE that simultaneously guide their visual attention. Interestingly the problem-solving success of experts also increased significantly.

By analyzing the visual attention we could conclude that the increased problem-solving success is accompanied by a shift in visual attention especially for novices between the non-MR and the MR condition in the first task of understanding the circuit. While in the non-MR group, novices tended to focus strongly on the middle branch of the circuit, in the MR group, attention also significantly shifted to the elements in the upper branch of the circuit. In contrast to the non-MR condition, in the MRLE the novices' attention is drawn to all relevant parts of the experiment. Furthermore, increased gaze patterns along the wires can be detected in the MRLE. Based on the analysis of visual attention, we can surmise which visualizations helped the participants in the MRLE. The increased gaze patterns along the wires in both groups were presumably caused by the visualization of the electron flow. We suspect that this has provided a better overview and global understanding of the circuit. Furthermore, the electron flow gives direct feedback where current is flowing and where not. Together with the charge level indication of the capacitors the strong focus of the

novices on the lower branch was probably reduced to a certain extent.

V. CONCLUSION

In this study we evaluated the use of MRLE in experimental problem solving and found that it induces a significant and large improvement in problem-solving success for both novices and experts. This partial result already suggests that the use of MRLE is promising in educational contexts, especially for experiment-related learning environments such as lab courses or similar. The use of MRLE might also be beneficial in other disciplines or contexts. Further research in these directions is needed. However, the use of MR-HMD in the classroom still poses a major challenge as it requires high-level hardware and IT infrastructure, technical expertise and suitable pedagogical framework conditions. Thus it seems currently only to be feasible in small study groups and rather on university level than for use in schools.

Furthermore we demonstrated that the use of MRLE directs the visual attention especially of novices to a more balanced and global view of the experiment rather than

being focused only on a few elements of the circuit. We also suggest that MRLE should not only be evaluated for their learning success, but that eye-tracking data should also be evaluated to gain insight into effects of design considerations and visualizations and to link the results back into further development of MRLE.

Future research should investigate whether the shift in visual attention is directly related to improved problem-solving. The use of any type of gaze guidance in educational MR settings opens a wide field of innovative and promising research.

ACKNOWLEDGMENTS

We acknowledge support by the Open Access Publication Funds of Technische Universität Braunschweig. This project is part of the “Qualitätsoffensive Lehrerbildung”, a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The programme is funded by the Federal Ministry of Education and Research. The authors are responsible for the content of this publication.

-
- [1] M. Strzys, S. Kapp, M. Thees, J. Kuhn, P. Lukowicz, P. Knierim, and A. Schmidt, Augmenting the thermal flux experiment: A mixed reality approach with the HoloLens, *Phys. Teach.* **55**, 376 (2017).
 - [2] O. Bodensiek, D. Sonntag, N. Wendorff, G. Albuquerque, and M. Magnor, Augmenting the fine beam tube: From hybrid measurements to magnetic field visualization, *Phys. Teach.* **57**, 262 (2019).
 - [3] N. D. Finkelstein, W. K. Adams, C. J. Keller, P. B. Kohl, K. K. Perkins, N. S. Podolefsky, S. Reid, and R. LeMaster, When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment, *Phys. Rev. ST Phys. Educ. Res.* **1**, 010103 (2005).
 - [4] C. E. Wieman, W. K. Adams, P. Loeblein, and K. K. Perkins, Teaching Physics Using PhET Simulations, *Phys. Teach.* **48**, 225 (2010).
 - [5] D. Sonntag, G. Albuquerque, M. Magnor, and O. Bodensiek, Hybrid learning environments by data-driven augmented reality, *Procedia Manuf.* **31**, 32 (2019).
 - [6] M. Strzys, S. Kapp, M. Thees, P. Klein, P. Lukowicz, P. Knierim, A. Schmidt, and J. Kuhn, Physics holo. Lab learning experience: Using smartglasses for augmented reality labwork to foster the concepts of heat conduction, *Eur. J. Phys.* **39**, 035703 (2018).
 - [7] D. Scaravetti and D. Doroszewski, Augmented reality experiment in higher education, for complex system appropriation in mechanical design, *Procedia CIRP.* **84**, 197 (2019).
 - [8] G. Westerfield, A. Mitrovic, and M. Billingham, Intelligent augmented reality training for assembly tasks, *Artificial Intelligence In Education* **7926**, 542 (2013).
 - [9] B. Volmer, J. Baumeister, S. Von Itzstein, I. Bornkessel-Schlesewsky, M. Schlewsky, M. Billingham, and B. Thomas, A comparison of predictive spatial augmented reality cues for procedural tasks, *IEEE Trans. Visualization Comput. Graphics*, **24**, 2846 (2018).
 - [10] N. Rappa, S. Ledger, T. Teo, K. Wai Wong, B. Power, and B. Hilliard, The use of eye tracking technology to explore learning and performance within virtual reality and mixed reality settings: A scoping review, *Interact. Learn. Environ.* **1** (2019).
 - [11] S. Kapp, M. Barz, S. Mukhametov, D. Sonntag, and J. Kuhn, ARETT: Augmented reality eye tracking toolkit for head mounted displays, *Sensors* **21** (2021).
 - [12] A. Poole and L. Ball, Eye tracking in human-computer interaction and usability research: Current status and future prospects, *Encyclopedia of Human Computer Interaction* (Idea Group Reference, UK, 2006), pp. 211–219.
 - [13] B. Ibrahim and L. Ding, Sequential and simultaneous synthesis problem solving: A comparison of students’ gaze transitions., *Phys. Rev. Phys. Educ. Res.* **17**, 010126 (2021).
 - [14] C. Wu and C. Liu, Eye-movement study of high- and low-prior-knowledge students’ scientific argumentations with multiple representations, *Phys. Rev. Phys. Educ. Res.* **17**, 010125 (2021).

- [15] E. Ince, An overview of problem solving studies in physics education, *J. Educ. Learning* **7**, 191 (2018).
- [16] M. Chi, P. Feltovich, and R. Glaser, Categorization and representation of physics problems by experts and novices, *Cogn. Sci.* **5**, 121 (1981).
- [17] P. Kohl and N. Finkelstein, Patterns of multiple representation use by experts and novices during physics problem solving, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010111 (2008).
- [18] G. Pólya, *How to Solve It: A New Aspect of Mathematical Method* (Princeton University Press, Princeton, NJ, 1985).
- [19] A. Newell and H. Simon, *Others Human Problem Solving* (Prentice-Hall, Englewood Cliffs, NJ, 1972).
- [20] J. Larkin, The role of problem representation in physics, *Mental Models* 75 (1983).
- [21] A. Deshpande and I. Kim, The effects of augmented reality on improving spatial problem solving for object assembly, *Adv. Engin. Informatics*, **38**, 760 (2018).
- [22] D. Karagozlu, Determination of the impact of augmented reality application on the success and problem-solving skills of students, *Quality Quantity*, **52**, 2393 (2018).
- [23] D. Sonntag and O. Bodensiek, Eye-tracking-based design of mixed reality learning environments in STEM, *Proceedings of the INTED2021*, 15th International Technology, Education and Development Conference (IATED, 2021), pp. 9519–9524, [10.21125/inted.2021.1990](https://doi.org/10.21125/inted.2021.1990).
- [24] We report only the corrected significance based on Bonferroni throughout the paper.
- [25] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Lawrence Erlbaum Associates, Hillside, NJ, 1988).
- [26] H. Levene, *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (Stanford University Press, Palo Alto, 1960), pp. 278–292.
- [27] C. Huberty and S. Olejnik, *Applied MANOVA and Discriminant Analysis* (John Wiley & Sons, New York, 2006).