E-ISSN: 3025-0994
Halaman 3474-3483
Volume 4 Nomor 4 Tahun 2026

From Theory to Spreadsheet: Evaluating the Pedagogical Effectiveness of Excel Solver in Teaching Modern Portfolio Theory

Hasnidar[1], Muhammad Zaikin[2]

[1,2] Department of Accounting, Universitas Halu Oleo
e-mail: hasnidar@uho.ac.id

Abstract

Spreadsheet optimisers have made the Markowitz model easy to compute in class, but computing a portfolio is not the same as understanding it. This study asked how far Excel Solver supports learning of modern portfolio theory and whether the design of a Solver task changes what students learn. Three dimensions were evaluated, namely procedural accuracy, construction of the efficient frontier, and reasoning about Solver output. Using qualitative document analysis, the study examined practicum reports written by student groups in two accounting classes. In the first class, groups set their own risk targets, compared Solver portfolios with randomly weighted ones and chose a best portfolio. In the second, groups traced fifteen efficient portfolios and drew the frontier. Every report was coded on the three dimensions, and its key figures were reproduced from the students' own price data. Solver performed reliably as a calculator, but its value for learning fell sharply when students had to place the frontier correctly and interpret the results. The two designs helped in complementary ways. Drawing the frontier led students to anchor their targets at the lowest-risk portfolio, whereas random portfolios supplied a ready benchmark that invited dominance comparisons. Weights that Solver drove to zero were the strongest prompt for conceptual explanation. Solver therefore becomes a learning tool rather than a black box only when the task obliges students to begin at the minimum-variance portfolio, benchmark the output and justify every weight.

Keywords: *Excel Solver, Modern Portfolio Theory, Efficient Frontier, Accounting Education, Task Design*

INTRODUCTION

Spreadsheets have become the everyday laboratory of quantitative finance teaching. They let students change an assumption and see its consequences at once, which turns models that would otherwise stay on the page into objects that can be explored, and simulation-based experiential learning has been associated with higher satisfaction among finance students (Bakoush, 2022). The same skills carry over into work, as structured Excel training has been shown to raise the competence and work readiness of finance students (Kirsten et al., 2025). Portfolio theory is an obvious candidate for this kind of teaching,

because its central calculation involves a matrix of covariances that is tedious by hand but immediate in a worksheet. For accounting students, who will later be asked to assess investment decisions for clients and employers, the spreadsheet is at once a learning device and a professional tool.

Markowitz (1952) showed that combining securities changes risk in a way that simple averaging cannot capture. Expected portfolio return is a weighted sum of security returns, but portfolio variance also contains every pairwise covariance, so securities that are not perfectly correlated offset part of each other's fluctuations. A rational investor should consider only portfolios for which no alternative offers more return for the same risk or less risk for the same return. Without short selling, this set, the efficient frontier, runs from the minimum variance portfolio up to the single security with the highest expected return (Elton et al., 2014). For a spreadsheet optimiser the implication is direct. A target-risk optimisation lands on the frontier only when the chosen risk level sits inside that interval, and any target outside it returns a portfolio that some other combination dominates.

Indonesian investment textbooks present the model as a sequence of calculations (Tandelilin, 2010; Hartono, 2017), and lecturers increasingly hand those calculations to Excel and its Solver add-in. Building accounting spreadsheets has been found to strengthen computational thinking (Muchsini et al., 2023), so the move to Excel has a sound pedagogical basis. Solver differs from an ordinary formula, however. Instead of deriving the weights, students specify a target and accept the algorithm's answer, so a task can be completed correctly without any grasp of why that answer is optimal.

This problem can be read through the distinction between procedural and conceptual knowledge. Procedural knowledge is the ability to carry out a sequence of steps, whereas conceptual knowledge is an understanding of the relationships behind those steps. The two develop iteratively and support each other, but procedural progress does not automatically produce conceptual understanding (Rittle-Johnson et al., 2001). Automated tools widen this gap. People tend to trust the output of automated systems too much and rarely check it (Parasuraman & Manzey, 2010), and a systematic review spanning several professional fields found this automation bias, the substitution of automated advice for vigilant checking, to be both common and consequential (Goddard et al., 2012). From the finance side, Michaud (1989) warned that mean-variance optimisers are highly sensitive to their inputs, and DeMiguel et al. (2009) showed that estimates from short samples are very unstable. Solver run correctly can therefore still lead to wrong conclusions when its targets and inputs are not understood.

Whether a tool supports conceptual learning depends partly on the task in which it is embedded. A Solver task can ask students to compare optimised portfolios with portfolios they weight themselves, or it can ask them to trace the efficient frontier point by point. Studies of spreadsheet use in accounting and finance education usually measure technical skills, confidence or perceptions, and few compare how different task designs shape what students actually learn. The evidence is available in the reports students write, which record how they set risk targets, read the curve, explain weights and test results. Analysis of such work products has been used before to map variation in accounting students' understanding of core disciplinary concepts (Lucas & Mladenovic, 2009).

This study therefore has two aims. The first is to evaluate the pedagogical effectiveness of Excel Solver in teaching modern portfolio theory along three dimensions, namely procedural accuracy, construction of the efficient frontier, and reasoning about the output. The second is to compare how two task designs used in two classes supported each of these dimensions. The results should help lecturers design Solver-based tasks that go beyond calculation skills.

METHOD

We adopted a qualitative evaluative approach in which written documents serve as the primary evidence (Bowen, 2009; Morgan, 2022). The corpus comprised every group report submitted in two parallel classes of the Capital Market and Portfolio Management course at the Department of Accounting, Faculty of Economics and Business, Universitas Halu Oleo, in October 2026, namely 32 reports prepared by 66 students. All groups worked with three securities and month-end closing prices from August 2025 to August 2026, although several stopped at July 2026. Their Excel templates used VAR for variances, COVAR for covariances and matrix multiplication for portfolio risk, and Solver was set to maximise return for a given risk with fully invested, non-negative weights. Identifying information was deleted before analysis, and the reports were relabelled K01 to K16 for Class A and C01 to C16 for Class D without reference to the original group numbers.

Each Class A group worked with a set of three stocks not used by any other group, generated four portfolios with arbitrary weights, ran Solver at eleven risk levels it selected itself, and nominated one portfolio as the best. Class D groups were free to include foreign securities and, starting from an initial allocation, produced fifteen Solver portfolios and drew the resulting frontier. Annual returns were obtained by multiplying the monthly mean by 12 in Class A and, in fifteen of the sixteen Class D groups, by compounding the monthly mean. Table 1 below contrasts the two task designs.

Table 1. Task designs in the two classes

Feature	Class A	Class D
Groups and students	16 groups, 33 students	16 groups, 33 students
Securities	Indonesian stocks, one US mix	Indonesian and foreign securities allowed
Portfolios required	4 random-weight and 11 Solver portfolios	Initial weights and 15 Solver portfolios
Risk targets	Chosen freely by the group	Chosen by the group to trace the frontier
Graph of the frontier	Not required	Required
Final output	Select the best portfolio	Describe the efficient portfolios
Annualisation of return	12 × mean monthly return	Compounded mean monthly return (15 groups)

Source: students' practicum reports, processed by the authors (2026)

Both templates implemented the textbook mean-variance expressions. With w_1, w_2 and w_3 denoting the weights of the three securities, portfolio expected return and variance are given by equations (1) and (2).

$$E(R_p) = w_1E(R_1) + w_2E(R_2) + w_3E(R_3) \quad (1)$$

$$\sigma_p^2 = \Sigma_i\Sigma_j\, w_iw_j\sigma_{ij} \quad (2)$$

In equation (2), σ_{ij} denotes the covariance of securities i and j, which becomes a variance when i equals j, and the square root of the result gives portfolio risk. Equation (3) shows the two ways in which the classes converted the mean monthly return $\bar{r}$ into an annual figure.

$$R_{annual} = 12 \times \bar{r} \text{ or } R_{annual} = (1 + \bar{r})^{12} - 1 \quad (3)$$

Coding moved through the condensation, display and verification stages described by Miles et al. (2014) and applied qualitative content analysis (Krippendorff, 2018). Each report was read end to end and coded against three evaluation dimensions. For procedural accuracy we checked whether the reported statistics were correct and whether each Solver point achieved the highest return available at its risk level. For frontier construction we located the chosen risk targets relative to the minimum-variance portfolio and the highest-return security. For reasoning we recorded how weights were explained, whether Solver results were checked against anything else, whether inputs were scrutinised, and what justified the final recommendation. On this basis each group was assigned to one of three levels of Solver use (Table 2), a classification that combines the procedural-conceptual distinction of Rittle-Johnson et al. (2001) with the SOLO principle of judging how ideas are connected rather than how many are listed (Biggs & Collis, 1982).

Table 2. Levels of Solver use in the students' reports

Level	What the report shows
Operational	Solver is run and its results are presented, but the weights and the shape of the curve are not explained or are misread. The efficient label is accepted without testing.
Constructive	Solver output is used to explain weights, the role of each stock or the shape of the frontier, but the frontier is still constructed incorrectly, the output is not tested, or the final decision reverts to a single stock or a dominated portfolio.
Reflective	The frontier is constructed correctly, the Solver output is tested against other portfolios or the plausibility of its inputs is questioned, and the final decision remains at the portfolio level.

Source: developed by the authors from Rittle-Johnson et al. (2001) and Biggs and Collis (1982)

All reports with usable price data were reproduced independently using the students' own conventions. A numerical optimiser without short positions traced each efficient frontier and located its minimum-variance portfolio, so that every Solver result could be set against the best return attainable at the same risk. This step served a diagnostic rather than an

evaluative purpose, separating arithmetic slips from conceptual misjudgements. Credibility was strengthened by cross-checking each report's tables, text and charts against the reproduced figures and by retaining the full coding sheet for audit.

RESULTS AND DISCUSSION

Table 3 lists the level of Solver use reached by each group, and Figure 1 compares the two classes. Across the 32 groups, nine reports were at the operational level, eleven at the constructive level and twelve at the reflective level. The classes differed mainly at the lowest level. Six groups in Class A stayed at the operational level, against three in Class D, while each class produced six reflective reports.

Table 3. Level of Solver use reached by each group

Code	Stocks	Level	Code	Stocks	Level
K01	BBCA, AADI, ANTM	Reflective	C01	SILO, INCO, ERAA	Constructive
K02	MDKA, ADMR, AMMN	Operational	C02	FORE, NSSS, BBTN	Reflective
K03	PTBA, ADRO, BUMI	Constructive	C03	ITMG, ICBP, BNGA	Reflective
K04	KAEF, ISAT, AKRA	Operational	C04	MAPI, MPMX, PTBA	Reflective
K05	NASA, APEX, ARTA	Reflective	C05	RACE, TSLA, AAPL	Operational
K06	ASII, ASSA, UNTR	Constructive	C06	BBCA, KJEN, BMRI	Constructive
K07	BELI, TINS, BBHI	Operational	C07	UNTR, INKP, AKRA	Constructive
K08	BBNI, ELSA, MRAT	Operational	C08	TLKM, OILS, ADES	Constructive
K09	BBTN, MPPA, MTEL	Reflective	C09	JPFA, SIDO, EXCL	Reflective
K10	BRPT, ITMG, INCO	Constructive	C10	ANTM, ISAT, CPIN	Constructive
K11	LPPF, CINT, ULTJ	Reflective	C11	BBNI, AMRT, GGRM	Constructive
K12	INKP, EXCL, GIAA	Reflective	C12	AADI, ADMR, INDF	Operational
K13	ICBP, MEDC, GGRM	Reflective	C13	ANTM, PGAS, ISAT	Reflective
K14	UDNG, JPFA, MEGA	Operational	C14	ASII, ICBP, ADRO	Operational
K15	WDC, TLKM, KO	Operational	C15	MYOR, ANTM, NVDA	Constructive
K16	TPIA, MBAP, PWON	Constructive	C16	BRIS, MSFT, 005930	Reflective

Source: students' practicum reports, coded by the authors (2026)

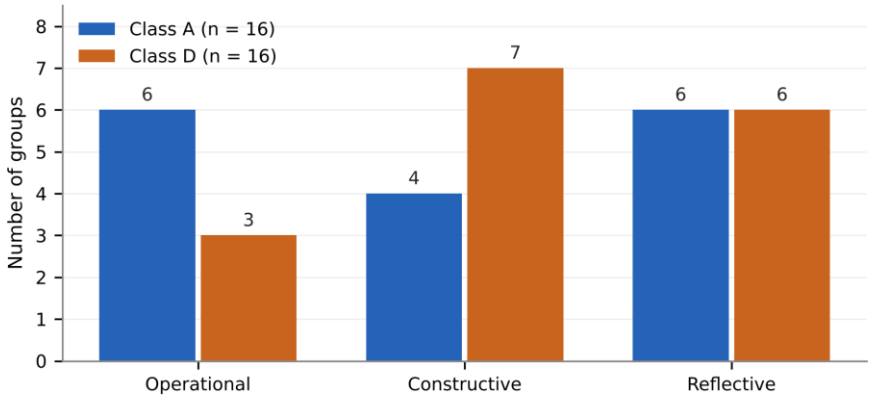


Figure 1. Levels of Solver use by class
Source: coding results, processed by the authors (2026)

On the dimension of procedural accuracy, Solver proved highly effective in both classes. Every report that supplied prices showed means, variances and standard deviations that our reproduction confirmed, while two groups in Class A showed only screenshots of their worksheets. In Class D, every Solver point checked lay within one percentage point of the maximum attainable return. In Class A, two groups obtained Solver results clearly below that maximum. K02's Solver output fell far below the frontier at several targets, so that one of its random portfolios outperformed most of its Solver portfolios, and K05's chosen Solver portfolio fell roughly two points short of the attainable maximum. The other errors in both classes arose in the parts done by the students themselves, such as weights summing to 101 percent, duplicated scenarios, copied figures that did not belong to the group's data, and target risks written in variance units. In other words, the algorithm was rarely wrong, and when it was, the cause lay in how it had been set up.

On the dimension of frontier construction, the effectiveness of Solver depended on the task design. In Class A, where groups chose their risk targets freely and no graph was required, two groups set targets outside the efficient segment. K16 placed all its targets above the volatility of its best-performing stock, so every Solver portfolio lay in the dominated region of the feasible set, and K06 extended its targets beyond the risk of its highest-return stock in the same way. Only two groups in Class A, K01 and K11, located the minimum-variance portfolio, and K01 was the only group in the class that justified its target range from theory, with the lower bound near the lowest attainable risk and the upper bound at the risk of the highest-return stock. In Class D, where the task asked for the frontier itself, every group started its target range at or above the minimum-variance portfolio, and six groups identified that portfolio correctly. One group in Class D, C15, still continued its targets beyond the highest-return stock, so that its last five points lay on the inefficient branch while being labelled efficient. Misidentification of the minimum-variance portfolio occurred once in each class, when K10 and C08 gave that label to their first Solver point without minimising risk.

Plotting the frontier, in short, helped students find where the curve starts, but did not by itself teach them where it ends.

On the dimension of reasoning, the most productive trigger came from Solver output that looked odd, namely weights that fell to zero or shrank sharply. When Solver gave a stock no weight, students were prompted to look for an explanation. C13 explained that PGAS received no weight because its return was negative and it moved together with the other two stocks, and C04 gave a similar explanation for MPMX. C03 concluded that, because ICBP received no weight in 13 scenarios, its efficient portfolios effectively consisted of two stocks. In Class A, K09 traced how the share of MTEL, a loss-making stock, shrank from 23.75 to 1.94 percent with each increase in target risk and attributed its early presence to its negative co-movement with the other holdings, while K01 observed that ANTM was squeezed out towards the high-risk end because AADI offered similar volatility with a better return. Reasoning of this kind applies dominance stock by stock, which lies at the heart of Markowitz's argument, and they support the view that procedural results can trigger conceptual understanding when they raise a question that must be answered (Rittle-Johnson et al., 2001). In contrast, C05 and C14 described the risk-return relationship on their frontier as linear, even though the graphs they had drawn were clearly concave, showing that a graph becomes a learning tool only when its shape must be interpreted.

Testing of Solver output followed the task design closely. In Class A, the four random-weight portfolios gave each group a ready benchmark, and eight groups compared them with the Solver portfolios at equal risk. K09 went further and solved one extra point at the risk level of a random portfolio that fell beyond its chosen targets, and K05 concluded from such a comparison that its Solver results were not absolute optimal points. The benchmark was not always used, however. K07 left its random portfolios out of the comparison and declared a Solver portfolio the best although one random portfolio had a higher return-to-risk ratio. In Class D, where no benchmark was required, six groups tested their Solver output on their own initiative, for example with a cloud of random portfolios in C13 or comparison portfolios in C16 and C15, and only C04 re-optimised its results independently. The remaining groups in both classes accepted the efficient label as it came out of Solver, a pattern consistent with automation bias (Parasuraman & Manzey, 2010; Goddard et al., 2012). The random portfolios of Class A therefore functioned as a simple but effective prompt for dominance reasoning.

The inputs entered into Solver were questioned in eleven reports, four in Class A and seven in Class D. Questioning was more frequent in Class D, partly because compounding the mean monthly return produced annual returns far above the actual price change for its most volatile stocks, which made the figures harder to ignore. Michaud (1989) described the optimiser as a tool that magnifies estimation error, and in both classes Solver pushed weights towards the stocks whose estimated returns were highest. Groups that compared the annualised return with the actual price change, such as K05 and K13 in Class A and C03 in Class D, were also among the most reflective users of Solver.

The closing instruction of each task left a clear imprint on the recommendations. Asked to name the best portfolio, seven Class A groups relied on a performance ratio, usually without any risk-free benchmark and sometimes still increasing at the last point

tested, so that, as in K13, the best portfolio was merely where the group had stopped computing. Two other Class A groups abandoned portfolio logic altogether, recommending a single stock (K03) or a portfolio dominated by another of their own options (K04). Asked instead to describe the frontier, twelve Class D groups matched portfolios to conservative, moderate or aggressive investors. Such matching is sound in principle, since the preferred point on the frontier is a matter of risk appetite, yet three Class D groups applied the rule that a portfolio becomes attractive once its percentage return exceeds its percentage risk, a rule that portfolio theory does not support. Table 4 compares the indicators for both classes.

Table 4. Indicators of Solver effectiveness by class

Dimension	Indicator	Class A	Class D
Procedural accuracy	Solver points clearly below the maximum attainable return	2	0
Frontier construction	Minimum-variance portfolio identified correctly	2	6
	Targets extending outside the efficient segment	2	1
	Minimum-variance portfolio misidentified	1	1
Reasoning about output	Solver output compared with other portfolios at equal risk	8	6
	Independent re-optimisation of Solver output	0	1
	Plausibility of inputs questioned	4	7
	Final decision based on a performance ratio	7	0
	Final decision left to investor risk profiles	4	12
	Decision reverting to a single stock or a dominated portfolio	2	1

Source: students' practicum reports and recomputation by the authors (2026)

Overall, the pedagogical effectiveness of Solver declines from the procedural to the reasoning dimension, and the two task designs strengthened different parts of learning. Plotting the frontier helped students anchor their targets at the minimum-variance portfolio and kept Solver output on the efficient segment, while random-weight portfolios gave a built-in benchmark that prompted dominance comparisons. Asking for a single best portfolio encouraged the use of ratios but also exposed their misuse, whereas asking for a description of the frontier encouraged the language of risk profiles. This finding supports the view that spreadsheet skills are beneficial (Muchsini et al., 2023; Kirsten et al., 2025) but shows that the benefit reaches conceptual understanding only when the task demands it. What distinguished the reflective groups in both classes was not their skill in operating Solver but their habit of questioning the results they had produced.

These findings suggest combining the strengths of both designs. Students should first compute the minimum-variance portfolio and then set target risks no higher than the volatility of the top-returning security, plotting every Solver point on the frontier. A small set of random-weight portfolios should be kept as a benchmark, and the efficient label should be

used only after each of them has been compared with the frontier at equal risk. Every zero or sharply falling weight should be explained in writing, since this is the part that most often triggers conceptual reasoning. The final decision should state its criterion explicitly, and any performance ratio should be examined across the whole frontier rather than only within the tested range.

Several caveats apply. The evidence comes from two classes in a single institution and is not intended for statistical generalisation. Because the classes differed in more than one task feature, no individual feature can be credited with the contrasts observed. Effectiveness was inferred from finished work rather than from measurements before and after an intervention, so no causal claim is made. The reports were written collaboratively and may contain text produced with generative AI (Kasneci et al., 2023), which the documents themselves cannot reveal. The material was coded by the research team without an independent second coder, a weakness partly offset by the numerical reproduction of every key claim. A quasi-experiment testing the combined design proposed here, supplemented by individual interviews, would be a natural next step.

CONCLUSION

Excel Solver is effective as a calculation tool in teaching modern portfolio theory, but its effectiveness declines for frontier construction and is lowest for reasoning. Plotting the frontier helped students locate its starting point, random portfolios prompted dominance comparisons, and zero weights triggered the most conceptual reasoning. Lecturers are advised to combine both designs by requiring the minimum-variance portfolio, a plotted frontier, random benchmark portfolios, and a written explanation of every zero weight.

ACKNOWLEDGEMENTS

We thank the 2026 Capital Market and Portfolio Management students in the Department of Accounting, Universitas Halu Oleo, whose practicum work made this study possible.

REFERENCES

Bakoush, M. 2022. Evaluating The Role Of Simulation-Based Experiential Learning In Improving Satisfaction Of Finance Students. *The International Journal Of Management Education*, 20(3): 100690. Https://Doi.Org/10.1016/J.Ijme.2022.100690

Biggs, J.B. & Collis, K.F. 1982. *Evaluating The Quality Of Learning: The SOLO Taxonomy*. New York: Academic Press.

Bowen, G.A. 2009. Document Analysis As A Qualitative Research Method. *Qualitative Research Journal*, 9(2): 27-40. Https://Doi.Org/10.3316/QRJ0902027

Demiguel, V., Garlappi, L. & Uppal, R. 2009. Optimal Versus Naive Diversification: How Inefficient Is The 1/N Portfolio Strategy? *Review Of Financial Studies*, 22(5): 1915-1953. Https://Doi.Org/10.1093/Rfs/Hhm075

Elton, E.J., Gruber, M.J., Brown, S.J. & Goetzmann, W.N. 2014. *Modern Portfolio Theory And Investment Analysis* (9th Ed.). Hoboken: John Wiley & Sons.

Goddard, K., Roudsari, A. & Wyatt, J.C. 2012. Automation Bias: A Systematic Review Of Frequency, Effect Mediators, And Mitigators. *Journal Of The American Medical Informatics Association*, 19(1): 121-127. Https://Doi.Org/10.1136/Amiajnl-2011-000089

Hartono, J. 2017. *Teori Portofolio Dan Analisis Investasi* (11th Ed.). Yogyakarta: BPFE.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J. & Kasneci, G. 2023. Chatgpt For Good? On Opportunities And Challenges Of Large Language Models For Education. *Learning And Individual Differences*, 103: 102274. Https://Doi.Org/10.1016/J.Lindif.2023.102274

Kirsten, E., Van Hoepen, R. & Du Toit, E. 2025. Can MS Excel Help Finance Students To Excel? A Study In Student Work Readiness. *South African Journal Of Higher Education*, 39(3): 155-172. Https://Doi.Org/10.20853/39-3-6383

Krippendorff, K. 2018. *Content Analysis: An Introduction To Its Methodology* (4th Ed.). Thousand Oaks: SAGE Publications.

Lucas, U. & Mladenovic, R. 2009. The Identification Of Variation In Students' Understandings Of Disciplinary Concepts: The Application Of The SOLO Taxonomy Within Introductory Accounting. *Higher Education*, 58(2): 257-283. Https://Doi.Org/10.1007/S10734-009-9218-9

Markowitz, H. 1952. Portfolio Selection. *The Journal Of Finance*, 7(1): 77-91. Https://Doi.Org/10.1111/J.1540-6261.1952.Tb01525.X

Michaud, R.O. 1989. The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal*, 45(1): 31-42. Https://Doi.Org/10.2469/Faj.V45.N1.31

Miles, M.B., Huberman, A.M. & Saldaña, J. 2014. *Qualitative Data Analysis: A Methods Sourcebook* (3rd Ed.). Thousand Oaks: SAGE Publications.

Morgan, H. 2022. Conducting A Qualitative Document Analysis. *The Qualitative Report*, 27(1): 64-77. Https://Doi.Org/10.46743/2160-3715/2022.5044

Muchsini, B., Siswandari, S., Gunarhadi, G. & Wiranto, W. 2023. Promoting College Students' Computational Thinking: The Use Of Constructionism-Based Accounting Spreadsheets Designing Activities. *Cogent Education*, 10(1): 2222866. Https://Doi.Org/10.1080/2331186X.2023.2222866

Parasuraman, R. & Manzey, D.H. 2010. Complacency And Bias In Human Use Of Automation: An Attentional Integration. *Human Factors*, 52(3): 381-410. Https://Doi.Org/10.1177/0018720810376055

Rittle-Johnson, B., Siegler, R.S. & Alibali, M.W. 2001. Developing Conceptual Understanding And Procedural Skill In Mathematics: An Iterative Process. *Journal Of Educational Psychology*, 93(2): 346-362. Https://Doi.Org/10.1037/0022-0663.93.2.346

Tandelilin, E. 2010. *Portofolio Dan Investasi: Teori Dan Aplikasi*. Yogyakarta: Kanisius.