Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Comparison of Experts and Novices in Determining the Gravitational Acceleration using Mobile Phone with Phyphox Application

Aungtinee KITTIRAVECHOTE* & Thanida SUJARITTHAM
Program of General Science, Faculty of Education, Bansomdejchaopraya Rajabhat University, Bangkok 10600, Thailand
*aungtinee.ki@bsru.ac.th

Abstract: With the advancement of technology, mobile phones have become the preferred instrument for monitoring a variety of phenomena in the science classroom. There are numerous applications designed to measure the magnitude of gravitational acceleration, Phyphox being one that has yet to be tested. In this study, we comprehensively assess the reliability of Phyphox application through 35 pre-service students (novices) and 5 experts by investigating the experimental mean gravity and limitations and response, specifically the differences between novices and experts in determining gravitational acceleration using a mobile phone with Phyphox and three mathematical analyses: taking average after formula substitution, manual plot, and commercial plot. The results showed that the experimental values of gravitational acceleration done by novices and experts match the theoretical value reported by the National Institute of Metrology (Thailand), with no significant difference in performance between them at .10 level. However, when comparing novices' performance across three mathematical analyses, novices' responses for manually plot differ significantly from those obtained for taking average after formula substitution and commercial plot at .05 level, which is due to some constraints related to novice plotting skills and the minimum experiment height of 0.100 m. This work paves the way for pre-service teachers to use Phyphox to determine the magnitude of gravitational acceleration using three mathematical analyses ranging from the most basic (i.e. taking average) to the most advanced (i.e. commercial plot) in order to assist them in assigning the free-falling experiment to schoolchildren using the most appropriate method of analysis.

Keywords: Phyphox, mobile phone experiment, experimental tools, gravitational acceleration

1. Introduction

The gravitational field strength is simply measured in terms of gravitational acceleration, which is a parabolic function of both displacement and time (Tate, 1968; White et al., 2007). Accurate estimation of its magnitude at various heights is critical for gaining an advantage in transitioning from school (Wang et al., 2017) to military (Chakravarthi, 2011) or astronomy work (Micheli et al., 2018). For example, using a ticker timer to measure the acceleration of gravity is considered to be a typical way in physics classrooms to highlight schoolchildren' potential in scientific data analysis (Fontana, Yeung & Hall, 2020), although measuring gravity with an absolute gravimeter is frequent at either a national or international institution with an accuracy of one in a billion parts (Ménoret et al., 2018). The desire for high-resolution gravity data, technological advancements in device downsizing, and a growing desire for knowledge and information, all have the potential to move the paradigm away from traditional fixed locations and toward mobile monitoring (Widiatmoko, Srigutomo & Kurniasih, 2012). Consequently, to quantify and estimate the magnitude of the gravitational acceleration, mobile sensing and monitoring devices with potential applications such as mobile phones are emerging (Kittiravechote & Sujarittham, 2020; Kuhn, 2014; Kuhn & Vogt, 2013; Pili, Violanda & Ceniza, 2018).

Because a mobile phone operates with a sensor known as micro-electro-mechanical systems, it has been an understandable choice for rapid laboratory equipment to report the strength of gravity (Kuhn & Vogt, 2013). Moreover, several studies use a microphone port or a magnetic field sensor to collect and compute the necessary data, such as the change in sound frequency (Kuhn, 2014) or the period of oscillation (Pili, Violanda & Ceniza, 2018), which leads to the determination of gravitational acceleration. More recently, we have published a paper that used a mobile phone with Phyphox application operating in timer mode and an acoustic stopwatch function to gauge the time of two acoustic events that were started and terminated by two sounds consecutively, and then evaluated the magnitude of the gravitational acceleration (Kittiravechote & Sujarittham, 2020). According to our findings, the experimental result for gravitational strength is determined nearly the standard value of 9.783 m/s^2 provided by the National Institute of Metrology (Thailand). Towards this end, all of these findings emphasize the advantages of using a mobile phone to measure the gravitational acceleration.

In this work, we are interested to address whether there is a difference between novices and experts in determining the acceleration of gravity using a mobile phone with Phyphox application and three mathematical analyses: taking average after formula substitution, manual plot, and commercial plot. To address the question, we utilize the inference statistic t-test to compare the responses of novices and experts: Welch's t-test is used to find the difference between novices and experts, while the paired sample t-test is used to illustrate the difference on three mathematical analyses done by novices. As such mathematical analysis is one of many ways for schoolchildren to interpret scientific data using methods ranging from the most basic (i.e. taking average) to the most advanced (i.e. commercial plot), this study suggested future work to help pre-service teachers in understanding where schoolchildren's difficulties originate during experimentation, resulting in the selection of the most appropriate method of analysis for schoolchildren at various levels.

2. Experiment

2.1 Mobile Phone Application

This work makes use of Phyphox application (physical phone experiment, version 1.1.7), a free program with at least 30 functions for conducting physics experiments (Staacks et al., 2018). The acoustic stopwatch feature is used as a timer to get the time between two loud acoustic signals. Clicks, beeps, claps, and other sounds are allowed as long as they are louder than ambient noise. The threshold is simply set to 0.3 arbitrary units (from 0 to 1), indicating the level at which the stopwatch is triggered. After this feature is enabled, the stopwatch starts with the first sound that exceeds the threshold and ends with the second sound. To repeat the experiment, simply delete the data and start over. It is important to note that the first sound is brief because a long sound could be misinterpreted as a pause.

2.2 Method

The procedure shown in figure 1 has previously mentioned for conducting the experiment (Kittiravechote & Sujarittham, 2020). In brief, A4 papers taped and folded into a design capable of holding a pencil at a height of approximately 0.100–1.500 meters are used, as illustrated in figure 1 (a). Phyphox application with a timer mode and an acoustic stopwatch feature is used to measure the time spent free-falling, as presented in figure 1 (b). A quick flick of the paper activates Phyphox application and displays the time on the screen, as shown in figure 1 (c). The stopwatch starts and stops timing in response to the first sound of the paper flicking and the second sound of the pencil hitting the floor. After recording and analyzing the results between height in meters and time in seconds, the magnitude of gravitational acceleration is calculated. It should be noted that this experiment is carried out at five different heights, each of which was replicated three times.

The data analysis is conducted by doing as described by Kittiravechote and Sujarittham (2020). In a nutshell, novices and experts are given the task of reporting the averaged free-fall time at each of the five heights, allowing the magnitude of the gravitational acceleration (g) to be determined by three methods: 1) taking the average of acceleration at all five different heights after substituting the averaged time (t) and height (h) into the formula $g = 2h/t^2$, 2) manual plot the linear graph $h = gt^2/2$ when h and t^2 present the data along the y and x axes, respectively, and the twice of its slope becomes the magnitude of gravity, and 3) using a commercial application like Microsoft Excel to map the straight line and estimate the magnitude of gravity (double the slope) (Meyer & Avery, 2009; Ose, 2016).

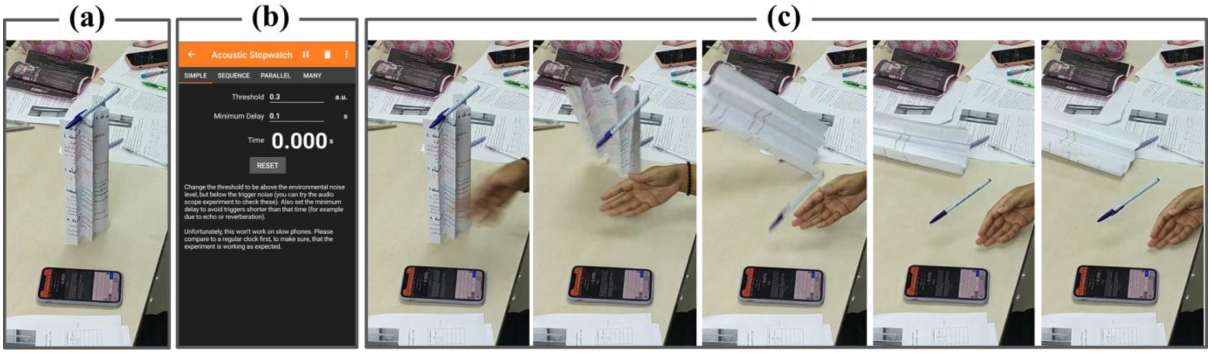

Figure 1. Experimental Method. (a) Setting the object's height. (b) Phyphox application with a timer mode and an acoustic stopwatch feature to measure the time spent free-falling. (c) Time series images capture the falling event.

2.3 Participants

35 pre-service teachers (novices) from Program of General Science, Faculty of Education, Bansomdejchaopraya Rajabhat University and 5 physics or general science lecturers (experts) were participated in this work. They all worked independently.

2.4 Statistical Inferences

Welch's t-test was used to assess the statistical significance of the difference in average changes between novices and experts. Paired sample t-test was used to determine the significance of the difference in average changes between the novices' mathematical analyses.

3. Result

The determination of the magnitude of gravitational acceleration (in m/s^2) obtained by 35 novices and 5 experts are displayed in figure 2. As shown in figure 2 (a) and (b), the mean and standard deviation of the magnitude of the gravitational acceleration for novices, g (taking average) = 9.469 ± 1.057, g (manual plot) = 10.120 ± 1.978, and g (commercial plot) = 9.452 ± 0.897, and for experts, g (taking average) = 9.618 ± 0.315, g (manual plot) = 9.691 ± 0.192, and g (commercial plot) = 9.701 ± 0.223. The results indicate that the mean values of both groups are comparable, e.g., g (taking average): 9.469 for novices vs. 9.618 for experts, g (manual plot): 10.120 for novices vs. 9.691 for experts, and g (commercial plot): 9.452 for novices vs. 9.701 for experts. The difference between the two means g by taking average, manual plot, and commercial plot are 0.149, 0.429, and 0.249 m/s^2, respectively, a marginal difference, suggesting that laboratory studies with novices and experts achieve nearly identical positive effects on the magnitude of gravitational acceleration. The straight lines connecting the three data points are also noted that they were all obtained during the same experiment with the same participant. In addition, the data distributions represented by boxplots in figure 2 (c) for novices and (d) for experts indicate that they are not only free of outliers but also nearly symmetrical, or in other words, the shapes are not overly skewed. Therefore, the possibility of using Welch's t-test to determine the significance of the differences in average changes between novices and experts is now available.

Thanks to the robustness of Welch's t-test for a comparison of two means with unequal sample sizes (i.e. 5 and 100) and unequal variances (i.e. 1 and 2) under normality as described by Derrick, Toher and White (2016), and thus the degrees of freedom of Welch's t-test are a random variable based on the sample size and variance of each sample, allowing us to perform Welch's t-test for 35 novices and 5 experts. Consequently, the null hypothesis: mean(Novices) – mean(Experts) = 0, along with the alternative hypothesis: mean(Novices) – mean(Experts) ≠ 0 were then defined. Following the calculations, the results in table 1 showed that the t-values are 0.655 with 21 degrees of freedom (df), 1.243 with 37 df, and 1.372 with 27 df for taking average, manual plot, and commercial plot, respectively. According to the t-distribution table (two-tails), all of the obtained t-values for the .10

level are less than the specified values: 0.655 < 1.721 at 21 df, 1.243 < 1.687 at 37 df, and 1.372 < 1.703 at 27 df. As a result, the null hypothesis was accepted in its entirety, while the alternative hypothesis was rejected. In other words, the mean differences between novices and experts were negligible at the 10% level. We suggested that the findings from novices and experts are not different by using a mobile phone with phyphox application to time individual falling objects at each height, together with three mathematical analyses: taking average by substitution of formulas, manually plotting of graphs, and plotting graphs with a program in order to find the gravitational acceleration of the earth at Bangkok.

In addition, paired sample t-test was used to see whether the difference in average changes between the novices' mathematical methods in figure 2 (a) was significant. According to the findings in table 2, there is no significant difference in the magnitude of gravitational acceleration when novices used the methods of taking average after substitution into the formula or commercial plot to evaluate the magnitude as our t-value (0.200) is less than the t-value at .05 level (2.032). When the novices graph on it manually, however, there is a noticeable difference at the .05 level as our t-values (2.244 and 2.051) are higher than the specified ones. As a result, because of the significant difference at the .05 level, there might be a feature of unpredictability in novices' manual plot capabilities.

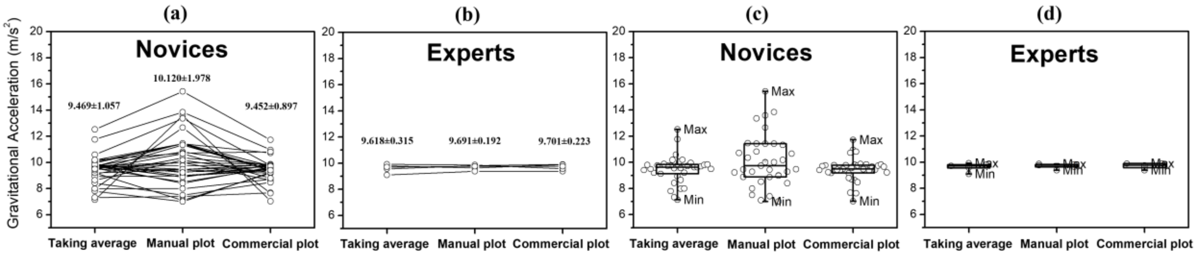

Figure 2. The Determination of the Magnitude of Gravitational Acceleration (in m/s^2). Measures of central tendency and spread (mean ± standard deviation) from novices (a) and experts (b). Nearly symmetric boxplots with no outliers from novices (c) and experts (d) provide the possibility of using Welch's t-test for significant difference.

Table 1. *Welch's T-Test for the Significant Differences in Performance between Novices and Experts*

	Taking average	Manual plot	Commercial plot
Welch's t-test	$t=\frac{\lvert 9.469-9.618\rvert-0}{\sqrt{\frac{1.057^2}{35}+\frac{0.315^2}{5}}}$ $t=0.655$ $df=21$	$t=\frac{\lvert 10.120-9.691\rvert-0}{\sqrt{\frac{1.978^2}{35}+\frac{0.192^2}{5}}}$ $t=1.243$ $df=37$	$t=\frac{\lvert 9.452-9.701\rvert-0}{\sqrt{\frac{0.897^2}{35}+\frac{0.223^2}{5}}}$ $t=1.372$ $df=27$
Noted:	$t=\frac{\lvert \overline{X}_{Novices}-\overline{X}_{Experts}\rvert-\lvert \overline{\mu}_{Novices}-\overline{\mu}_{Experts}\rvert}{\sqrt{\frac{S.D._{Novices}^2}{n_{Novices}}+\frac{S.D._{Experts}^2}{n_{Experts}}}}$	$df=\frac{\left(\frac{S.D._{Novices}^2}{n_{Novices}}+\frac{S.D._{Experts}^2}{n_{Experts}}\right)^2}{\frac{\left(\frac{S.D._{Novices}^2}{n_{Novices}}\right)^2}{n_{Novices}-1}+\frac{\left(\frac{S.D._{Experts}^2}{n_{Experts}}\right)^2}{n_{Experts}-1}}$	

Table 2. *Paired Sample T-Test for the Significant Differences between the Novices' Mathematical Methods*

Comparison	T-statistic with 34 df	Decision
Taking average vs. manual plot	2.244	Significantly different
Taking average vs. commercial plot	0.200	Not significantly different
Manual plot vs. commercial plot	2.051	Significantly different

Noted: T-statistic with 34 df at .05 level is 2.032

4. Discussion

In attempt to discover the gravitational acceleration of the earth at Bangkok, we have demonstrated how novices and experts use a mobile phone with Phyphox application to time individual falling objects at each height, as well as three mathematical analyses: taking average, manual plot, and commercial plot. At the .10 level, we found that novices and experts performed equally, suggesting that our method of measuring gravitational acceleration is practical and good enough for pre-service teachers to conduct classroom experiments. Further to that, for the novices with the use of three mathematical analyses, we found an insignificant difference between the methods of taking average after formula substitution and commercial plot, but not in the case of manual plot at the .05 level.

To validate the reliability of the magnitude of gravitational acceleration obtained from our experiment, we compare all means to the theoretical value of 9.783 m/s^2 proposed by the National Institute of Metrology (Thailand). All of the means (9.469, 10.120, 9.452, 9.618, 9.691, and 9.701 m/s^2) are comparable to the theoretical value. The maximum error is found to be 3.445 % when novices conduct the manual plot for data analysis. Moreover, the results of the Welch's t-test show that the difference between novices and experts in performing the experiment to determine the earth's gravitational acceleration is insignificant at 90% confidence. As a result, our experimental methods for measuring gravitational acceleration can produce reliable results.

As previously stated, the results of the pair sample t-test show significant differences in novices' data analysis using manual plots compared to taking average after formula substitution or commercial plots at 95 percent confidence interval, implying that there may be a variable that limits novices' manual plotting skills. We hence examine the novices' hand-drawn graphs and summarize the mistakes made by 22 out of 35 novices, as shown in table 3. Not surprisingly, these mistakes are identified as the most common graphical mistakes made by schoolchildren when it comes to slope conceptualization (Hattikudur et al., 2012). These findings suggested that pre-service teachers could identify and recognize faults in manual plot, allowing them to comprehend the problems that schoolchildren face with manual plot as well as the sources of errors or misconceptions.

In addition, we found that when novices performed the experiments at a height of less than 10 cm (about half the width of an A4 paper), some of them showed a misrepresentation of acceleration. They experienced the difficulties of a mobile phone in retrieving the falling time due to the use of a phone that was too late (which could be easily solved by changing the tool to the faster one). This also included Phyphox program's period time problem. However, since the object required 142 milliseconds to reach the ground at a height of 10.0 centimeters, Phyphox developer recommended that the acoustic stopwatch wait at least 100 milliseconds before accepting a second signal (AachenUniversity, November 2017). The results indicate limitations in our experiments, especially the need for a suitable height of greater than 0.100 m above the ground.

Table 3. *A Summary of the Mistakes Found in 22 of 35 Novices*

Description of mistakes	Novices
Incorrect scale. On the x-axis, y-axis, or both x and y-axes, the scale is incorrect. Novices should represent the same value on each axis, for example, if each division is set to 0.005 at the start, then the numbers along the axis should be 0.005, 0.010, 0.015, and so on. As previously discussed, this provides them to estimate the magnitude of gravity from the slope of best fit passing through the origin.	3
Unbalanced line of best fit. (1) Data points above and below the line of best fit appear to be less evenly distributed. (2) Draw a line that best describes the identified trend as a connection of data points, rather than a straight line. Novices should draw the linear line that best applies to the majority of the data points and should be less concerned with data points that differ from the majority. (3) Line of best fit does not pass through the origin.	22
Calculate slope from two specific data points. Using two raw data points to determine the graph's slope. To calculate the slope of a graph, novices could choose any two points that lie on the line of best fit.	2
Missing unit. On the x and y axes, the variables and units of measurement are missing. They can be identified by novices.	5

5. Conclusion

We have demonstrated the differences in how novices and experts conduct the experiment to determine the magnitude of the gravitational acceleration using a mobile phone with Phyphox application, as well as the data analysis with three mathematical methods: taking average after formula substitution, manual plot, and commercial plot. Not only do the responses of novices and experts match the theoretical value, but we also present a statistically negligible difference in their performance with a 90% confidence interval. Accordingly, our method of determining gravitational acceleration is reliable, making it possible for pre-service teachers doing classroom experiments with schoolchildren. Noted that the responses from novices after using the manual plot for data analysis differ significantly from those obtained by taking average and commercial plot with 95% confidence interval due to some limitations related to novice plotting skills and the minimum experiment height of 0.100 m.

Acknowledgements

Research and Development Institute, Bansomdejchaopraya Rajabhat University is acknowledged.

References

AachenUniversity, R. W. T. H. (November 2017). *Experiment: Acoustic Stopwatch*. https://phyphox.org/wiki/index.php/Experiment:_Acoustic_Stopwatch.

Chakravarthi, V. (2011). *Encyclopedia of Solid Earth Geophysics*. Dordrecht: Springer Netherlands.

Derrick, B., Toher, D., & White, P. (2016). Why Welch's test is Type I error robust. *Quantitative Methods for Psychology, 12*(1), 30-38.

Fontana, E., Yeung, C., & Hall, J. C. (2020). Determining the acceleration due to gravity and friction using the ticker tape timer method. *The Physics Teacher*, *58*(5), 338-339.

Hattikudur, S., Prather, R. W., Asquith, P., Alibali, M., Knuth, E. J., & Nathan, M. (2012). Constructing graphical representations: Middle schoolers' intuitions and developing knowledge about slope and y-intercept. *School Science and Mathematics*, *112*(4), 230-240.

Kittiravechote, A., & Sujarittham, T. (2020). Measuring the acceleration of gravity using a smartphone, A4-papers, and a pencil. *International Journal of Advanced Science and Technology*, *29*(7s), 884-889.

Kuhn, J. (2014). Relevant information about using a mobile phone acceleration sensor in physics experiments. *American Journal of Physics*, *82*(2), 94-94.

Kuhn, J., & Vogt, P. (2013). Smartphones as experimental tools: Different methods to determine the gravitational acceleration in classroom physics by using everyday devices. *European J of Physics Education*, *4*(1), 16-27.

Ménoret, V., Vermeulen, P., Le Moigne, N., Bonvalot, S., Bouyer, P., Landragin, A. et al. (2018). Gravity measurements below 10^{-9} g with a transportable absolute quantum gravimeter. *Scientific Reports*, *8*(1), 12300.

Meyer, D. Z., & Avery, L. M. (2009). Excel as a qualitative data analysis tool. *Field Methods*, *21*(1), 91-112.

Micheli, M., Farnocchia, D., Meech, K. J., Buie, M. W., Hainaut, O. R., Prialnik, D. et al. (2018). Non-gravitational acceleration in the trajectory of 1I/2017 U1 ('Oumuamua). *Nature*, *559*(7713), 223-226.

Ose, S. O. (2016). Using excel and word to structure qualitative data. *Journal of Applied Social Sciences*, *10*(2), 147-162.

Pili, U., Violanda, R., & Ceniza, C. (2018). Measurement of g using a magnetic pendulum and a smartphone magnetometer. *The Physics Teacher*, *56*(4), 258-259.

Staacks, S., Hütz, S., Heinke, H., & Stampfer, C. (2018). Advanced tools for smartphone-based experiments: phyphox. *Physics Education*, *53*(4), 045009.

Tate, D. R. (1968). Acceleration due to gravity at the National Bureau of Standards. *Journal of Research of the National Bureau of Standards, Section C: Engineering and Instrumentation*, *72C*(1), 1-20.

Wang, Q., Wang, C., Xiao, Y., Schulte, J., & Shi, Q. (2017). A new method of measuring gravitational acceleration in an undergraduate laboratory program. *European Journal of Physics*, *39*(1), 015701.

White, J. A., Medina, A., Román, F. L., & Velasco, S. (2007). A measurement of g listening to falling balls. *The Physics Teacher*, *45*(3), 175-177.

Widiatmoko, E., Srigutomo, W., & Kurniasih, N. (2012). Measurement of gravitational acceleration using a computer microphone port. *Physics Education*, *47*(6), 709-714.