Report on Assessment Study of Self-Instructional Materials Programme (SIMs)

A survey of Bhutanese SIM students, teachers, principals, district education officers, parents and community leaders

"Reaching The Unreached"

ཤེས་རིག

Department of School Education

Ministry of Education, Bhutan

Evidence Report
on
Nationwide SIM Assessment Study

Department of School Education
Ministry of Education, Bhutan

November 2021

TABLE OF CONTENTS

Executive Summary

Demographic characteristics of SIM survey respondents

1. **SIM students:** The age of the SIM student respondents ranged from 6 to 24 years (M = 13.38, SD = 3.50). Among the 2648 SIM student respondents, 1210 (45.7%) were males and 1438 (54.3%) were females. Likewise, among the 2648 SIM student respondents, we got data representation from all classes from Class I to Class XII with maximum from class VI (12.0%), closely followed by class X (11.4%), class VII (10.2%), class IX(10.2%), class V (9.7%), class IV (7.9%), class XII (7.4%), class III (6.7%), class XI (6.6%), class VIII (6.4%), class II (6.1%) and with minimum from class I (5.4%). Class PP students were not surveyed because they did not exist last year when SIM programme was implemented. Among the 2648 SIM student respondents, by school type also we got data representation from all types of schools with maximum from HSS (36.0%), followed by PS (30.7%), MSS (24.1%), LSS (6.7%), and with minimum from ECR (2.6%).

   SIM teachers: The age of the SIM teacher respondents ranged from 24 to 57 years (M = 33.85, SD = 6.45). Among the 667 SIM teacher respondents, 400 (60%) were males and 267 (40%) were females. Among the 667 SIM teacher respondents, we got data representation from all classes from Class PP to Class XII with maximum teaching class X (18.3%), followed by class XII (13.2%), class VI (12.1%), class I (7.8%), class III (6.5%), class IV (6.3%), class V (6.3%), class IX (6.3%), class VIII (6.0%), class VII (5.9%), class II (5.3%), class XI (6.1%) and minimum teaching class PP (1.2%). Among the 667 SIM teacher respondents, we got data representation from all types of schools such as HSS (43.5%), MSS (29.2%), LSS (7.8%), PS (18.7%), and ECR (0.8%).

   SIM principals: The age of the SIM principal respondents ranged from 28 to 65 years (M = 43.17, SD = 6.34). Among the 123 SIM principal respondents, 121 (98.4%) were males and 2 (1.6%) were females. Among the 123 SIM principal respondents, we got data representation from all types of schools such as HSS (18.7%), MSS (11.4%), LSS (7.3%), PS (57.7%), and ECR (4.9%).

   SIM DEOs: The age of the SIM DEO respondents ranged from 41 to 54 years (M = 48.24, SD = 4.09). Among the 29 SIM chief DEO and deputy DEO respondents, 26 (89.7%) were males and 3 (10.3%) were females.

   SIM LG leaders: The age of the SIM LG respondents ranged from 27 to 58 years (M = 37.67, SD = 6.82). Among the 76 SIM LG respondents, 65 (85.5%) were males and 11 (14.5%) were females.

   SIM parents: The age of the SIM parent respondents ranged from 19 to 72 years (M = 37.93, SD = 8.45). Among the 374 SIM principal respondents, 166 (44.4%) were males and 208 (55.6%) were females. Among the 374 SIM parent respondents, we got data representation from all types of schools such as HSS (15.2%), MSS (20.6%), LSS (11.5%), PS (40.4%), and ECR (12.3%). We also included question on special education needs (SEN) students. Among the 374 SIM parent respondents, 34 (9.1%) said their children are SEN students and 340 (90.9%) said their children are not SEN students.

Effectiveness of SIM Programme

2. **Satisfaction level of SIM programme**: The 74.4% of the SIM student respondents rated the SIM programme "satisfied" or "extremely satisfied" in our survey. Our survey also found that this is consistently same in all age groups, in all key stages and in all school types. Similarly, the 72.1% of the SIM teacher respondents rated the SIM programme "satisfied" or "extremely satisfied." The 87.0% of the SIM principal respondents rated the SIM programme "satisfied" or "extremely satisfied." The 89.6% of the SIM DEO respondents rated the SIM programme "satisfied" or "extremely satisfied." The 85.5% of the SIM LG leader respondents rated the SIM programme "satisfied" or "extremely satisfied."

   Evidence on SIM satisfaction level: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 74.4% of SIM students, both female students and male students, are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 25.537, p = 0.0000, with a moderate effect size (*r* = 0.50).

   Similarly, in the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority 72.1% of SIM teachers, both female teachers and male teachers, are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 11.830, p = 0.0000, with a moderate effect size (*r* = 0.46).

   In the SIM principal population, there is statistically significant evidence (p = 0.0000) that 87.0% of SIM principals are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.152, p = 0.0000, with a strong effect size (*r* = 0.74).

   In the SIM DEO population, there is statistically significant evidence (p = 0.0000) that 89.6% of SIM DEOs are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.186, p = 0.0000, with a strong effect size (*r* = 0.78).

   In the SIM LG leader population, there is statistically significant evidence (p = 0.0000) that 85.5% of SIM LG leaders are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample

Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.074, p = 0.0000, with a strong effect size (*r* = 0.70).

3. **Acceptance level of SIM programme**: The 72.1% of the SIM student respondents rated the SIM learning "enjoyable" or "extremely enjoyable" in our survey. Our survey also found that this is consistently same in all age groups, in all key stages and in all school types. However, only 35.8% of the SIM teacher respondents rated the SIM learning "enjoyable" or "extremely enjoyable" for their students. It means while acceptance level of SIM programme among SIM students was good, SIM teachers perceived that the acceptance level of SIM programme among their students was poor. Nevertheless, the 91.0% of the SIM principal respondents rated that the SIM programme "useful" or "very useful." The 93.1% of the SIM DEO respondents rated that the SIM programme "useful" or "very useful." The 82.9% of the SIM LG respondents rated that the SIM programme "useful" or "very useful." The 82.4% of the SIM parent respondents rated that the SIM programme "useful" or "very useful."

   Evidence on SIM acceptance level: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 72.1% of SIM students, both girls and boys, found SIM learning enjoyable during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 23.604, p = 0.0000, with a moderate effect size (*r* = 0.46).

   In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that only 35.8% of SIM teachers found SIM learning enjoyable during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly below hypothesized value of 2.5, Z = -6.949, p = 0.0000, with a low effect size (*r* = 0.27).

   In the SIM principal population, there is statistically significant evidence (p = 0.0000) that 91.0% of SIM principals believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.889, p = 0.0000, with a very strong effect size (*r* = 0.80).

   In the SIM DEO population, there is statistically significant evidence (p = 0.0000) that 93.1% of SIM DEOs believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.443, p = 0.0000, with a very strong effect size (*r* = 0.83).

   In the SIM LG leader population, there is statistically significant evidence (p = 0.0000) that 82.9% of SIM LG leaders believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 5.901, p = 0.0000, with a strong effect size (*r* = 0.68).

In the SIM parent population, there is statistically significant evidence ($p = 0.0000$) that 82.4% of SIM parents believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 12.518$, $p = 0.0000$, with a strong effect size ($r = 0.65$).

Effectiveness of SIM Materials

4. **Effectiveness of overall presentation of SIM materials:** The 81.0% of the SIM student respondents rated the overall presentation of SIM materials "effective" or "extremely effective" in our survey. Our survey also found that this is consistently same in all age groups and in all key stages. However, in school types, our data show that majority of school types such as HSS, MSS, LSS, and PS rated SIM overall presentation as "effective" while ECR rated SIM overall presentation as "extremely effective." Similarly, the 84.7% of the SIM teacher respondents rated the overall presentation of SIM materials "effective" or "extremely effective." Similarly, the 94.3% of the SIM principal respondents rated that overall presentation of SIM booklets is attractive. The 89.7% of the SIM DEO respondents rated that overall presentation of SIM booklets is attractive. The 93.4% of the SIM LG respondents rated that overall presentation of SIM booklets is attractive. The 93.6% of the SIM parent respondents rated that overall presentation of SIM booklets is attractive.

   Evidence on overall presentation of SIM materials: In the SIM student population, there is statistically significant evidence ($p = 0.0000$) that the majority 81.0% of SIM students, both girls and boys, found overall presentation of the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 32.003$, $p = 0.0000$, with a strong effect size ($r = 0.62$).

   In the SIM teacher population, there is statistically significant evidence ($p = 0.0000$) that the majority 84.7% of SIM teachers, both female teachers and male teachers, found overall presentation of the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 18.130$, $p = 0.0000$, with a strong effect size ($r = 0.70$).

   In the SIM principal population, there is statistically significant evidence ($p = 0.015409$) that at least 88% of SIM principals believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM principals who believe overall presentation of SIM is attractive ($N_{Yes} = 116$, 94.3%), was statistically significantly greater than the population hypothesized value of 88%, $p = 0.015409$.

   In the SIM DEO population, there is statistically significant evidence ($p = 0.0345460$) that at least 74% of SIM DEOs believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM DEOs who believe overall presentation of SIM is attractive ($N_{Yes} = 26$, 89.7%), was statistically significantly greater than the population hypothesized value of 74%, $p = 0.035460$.

   In the SIM LG leader population, there is statistically significant evidence ($p = 0.035814$) that at least 86% of SIM LG leaders believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM LG leaders who believe overall presentation of SIM is attractive (N_{Yes} = 71, 93.4%), was statistically significantly greater than the population hypothesized value of 86%, p = 0.035814.

In the SIM parent population, there is statistically significant evidence (p = 0.009820) that at least 90% of SIM parents believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM parents who believe overall presentation of SIM is attractive (N_{Yes} = 350, 93.6%), was statistically significantly greater than the population hypothesized value of 90%, p = 0.009820.

5. **Effectiveness of contents of SIM materials:** The 74.4% of the SIM student respondents rated the SIM contents "effective" or "extremely effective" in our survey. Our survey also found that this is consistently same in all age groups and in all key stages. However, in school types, our data show that majority of school types such as HSS, MSS, LSS, and PS rated SIM contents as "effective" while ECR rated SIM contents as "extremely effective." Similarly, the 78.1% of the SIM teacher respondents rated the SIM contents "effective" or "extremely effective."

   Evidence on contents of SIM materials: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 74.4% of SIM students, both girls and boys, found contents of SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 26.682, p = 0.0000, with a moderate effect size (r = 0.52).

   In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority 78.1% of SIM teachers, both female teachers and male teachers, found contents of SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.817, p = 0.0000, with a moderate effect size (r = 0.57).

6. **Effectiveness of instructions of SIM materials:** The 69.9% of the SIM student respondents rated the SIM instructions "effective" or "extremely effective" in our survey. Our survey also found that this is consistently same in all key stages and in all school types. However, in age groups, our data show that majority of age groups rated SIM instructions as "effective" except age group 20-24 which rated instructions as ineffective. But the difference is marginal and not significant. Similarly, the 77.2% of the SIM teacher respondents rated the SIM instructions "effective" or "extremely effective."

   Evidence on instructions of SIM materials: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 69.9% of SIM students, both girls and boys, found instructions in SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 22.345, p = 0.0000, with a moderate effect size (r = 0.43).

   In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority 77.2% of SIM teachers found instructions in SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.683, p = 0.0000, with a moderate effect size (r = 0.57).

7. **Effectiveness of graphics of SIM materials:** The 77.5% of the SIM student respondents rated the SIM graphics "effective" or "extremely effective" in our survey. Looking at students' rating of SIM graphics by age group, key stage and school type, it shows that older students, higher key stages or higher class level schools such as HSS, MSS and LSS rated SIM graphics as "effective" while younger children, lower key stages or lower class level schools such as ECR and PS rated SIM graphics as "extremely effective." This is an important and consistent finding. This will have an important policy implication for the future material designs of SIM booklets that it's more effective to include more graphics for lower classes. Similarly, the 81.1% of the SIM teacher respondents rated the SIM graphics "effective" or "extremely effective."

   Evidence on graphics of SIM materials: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 77.5% of SIM students, both girls and boys, found graphics in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 29.999, p = 0.0000, with a moderate effect size (r = 0.58).

   In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority 81.1% of SIM teachers, both female teachers and male teachers, found graphics in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 16.607, p = 0.0000, with a strong effect size (r = 0.64).

8. **Effectiveness of activities of SIM materials:** The 79.0% of the SIM student respondents rated the SIM activities "effective" or "extremely effective" in our survey. Our survey also found that this is consistently same in all age groups and in all key stages. However, in school types, our data show that majority of school types such as HSS, MSS, LSS, and PS rated SIM contents as "effective" while ECR rated SIM activities as "extremely effective." It seems lower classes appreciated activities more. Similarly, the 81.1% of the SIM teacher respondents rated the SIM activities "effective" or "extremely effective."

   Evidence on activities of SIM materials: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 79.0% of SIM students, both girls and boys, found activities in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 30.287, p = 0.0000, with a moderate effect size (r = 0.59).

   In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority 81.1% of SIM teachers, both female teachers and male teachers, found activities in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 16.395, p = 0.0000, with a strong effect size (r = 0.63).

Effectiveness of SIM Learning

9. **Effectiveness of SIM learning in increasing knowledge:** The 62.7% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in increasing their knowledge in comparison to classroom learning. Our survey also found that this is consistently same in all age groups, in all key stages, and in all school types. However, only 40.9% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in increasing knowledge.

   Evidence on effectiveness of SIM learning in increasing knowledge: In the SIM student population, there is statistically significant evidence ($p = 0.0000$) that the majority 62.7% of SIM students, both girls and boys, found SIM learning effective in increasing their knowledge. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 14.123$, $p = 0.0000$, with a low effect size ($r = 0.27$).

   In the SIM teacher population, there is statistically significant evidence ($p = 0.0000$) that only minority 40.9% of SIM teachers, both female teachers and male teachers, found SIM learning effective in increasing knowledge. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly below hypothesized value of 2.5, $Z = -5.063$, $p = 0.0000$, with a very low effect size ($r = 0.20$).

10. **Effectiveness of SIM learning in increasing skills:** The 56.9% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in increasing their skills in comparison to classroom learning. Looking at students' rating of SIM learning in increasing skills, by age group, key stage and school type, it shows that all age groups except 5-9 year old age group, all key stages except key stage I, and school types except ECR have rated SIM learning "effective" for increasing skills. Consistent with 5-9 year old age group and students in key stage I, ECR rated SIM learning "ineffective" for increasing skills. This clearly shows younger children struggled to learn skills during SIM learning. However, only 38.4% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in increasing skills.

    Evidence on effectiveness of SIM learning in increasing skills: In the SIM student population, there is statistically significant evidence ($p = 0.0000$) that the majority 56.9% of SIM students, with very low but significant difference between girls and boys, found SIM learning effective in increasing their skills. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 9.275$, $p = 0.0000$, with a very low effect size ($r = 0.18$).

    In the SIM teacher population, there is statistically significant evidence ($p = 0.0000$) that only minority 38.4% of SIM teachers, both female teachers and male teachers, found SIM learning effective in increasing skills. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = -6.202$, $p = 0.0000$, with a low effect size ($r = 0.24$).

11. **Effectiveness of SIM learning in imparting values:** The 54.6% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in imparting values in comparison to classroom learning. Our survey also found that this is consistently same in all age groups. However, for key stages and school types, the results were mixed. Majority of the key stages except key stages I and IV have rated SIM learning "effective" for imparting values. The key stages I and IV have rated it "ineffective." Similarly, majority of the school types have rated it "effective." But ECR and MSS have rated it "ineffective." However, only 29.0% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in imparting values.

    Evidence on effectiveness of SIM learning in imparting values: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 54.6% of SIM students found SIM learning effective in imparting values. However, there is a very low but significant difference between girls and boys where girls found SIM learning effective in imparting values but boys found it ineffective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.422, p = 0.0000, with a very low effect size (*r* = 0.13). The positive z-score shows that the population median is above the hypothesized median of 2.5.

    In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that only minority 29.0% of SIM teachers, both female teachers and males teachers, found SIM learning effective in imparting values. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly below hypothesized value of 2.5, Z = -11.121, p = 0.0000, with a moderate effect size (*r* = 0.43).

12. **Effectiveness of SIM learning in improving attitudes:** The 52.4% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in improving attitudes in comparison to classroom learning. Students' rating of SIM learning in improving attitudes by age group, key stage, and school type were mixed. The age groups 10-14 and 20-24 have rated SIM learning "effective" in improving attitudes. But the age groups 5-9 and 15-19 have rated it "ineffective." Similarly, the key stages II, III and V have rated it "effective". But the key stages I and IV have rated it "ineffective." Likewise, the majority of the school types have rated it "effective." However, ECR and MSS have rated it "ineffective." However, only 23.1% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in improving attitudes.

    Evidence on effectiveness of SIM learning in improving attitudes: In the SIM student population, there is statistically significant evidence (p = 0.0013) that the majority 52.4% of SIM students found SIM learning effective in improving attitudes. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 3.216, p = 0.0013, with a very low effect size (*r* = 0.06).

    In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that only minority 23.1% of SIM teachers, both female teachers and male teachers, found SIM learning effective in improving attitudes. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -14.332, p = 0.0000, with a moderate effect size (*r* = 0.56).

13. **Effectiveness of SIM learning in understanding English:** The 56.6% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in understanding English subject in comparison to classroom learning. Looking at students' rating of SIM learning in understanding English subject by age group, key stage, and school type, our data results show that the majority of the age groups except 5-9 age group, the majority of key stages except key stage I, and the majority of the school types except ECR have rated SIM learning "effective" in understanding English. But the age group 5-9, the key stage I, and ECR have rated it as "ineffective." It seems the younger children or students in lower classes had difficulty in understanding English during SIM learning. However, only 34.3% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in understanding English.

    Evidence on effectiveness of SIM learning in understanding English: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 56.6% of SIM students found SIM learning effective in understanding English subject. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.914, p = 0.0000, with a very low effect size (*r* = 0.17).

    In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that only minority 34.3% of SIM teachers, both female teachers and male teachers, found SIM learning effective in understanding English. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -8.331, p = 0.0000, with low effect size (*r* = 0.32).

14. **Effectiveness of SIM learning in understanding Mathematics:** Only 47.9% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in understanding Mathematics subject in comparison to classroom learning. Looking at students' rating of SIM learning in understanding Mathematics by age group, key stage, and school type, data results show that the majority of SIM students in all categories have rated Mathematics learning as "ineffective" during SIM learning. It seems the majority of the students had difficulty in understanding Mathematics during SIM learning. The sample median choice rating was 2, which is "ineffective." This means at least 50% of the SIM student respondents found SIM learning "ineffective" or "extremely ineffective" in understanding Mathematics. Similarly, only 20.6% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in understanding Mathematics.

    Evidence on effectiveness of SIM learning in understanding Mathematics: In the SIM student population, there is statistically significant evidence (p = 0.0002) that only minority 47.9% of SIM students found SIM learning effective in understanding Mathematics. In other words, the majority 52.1% of SIM students found SIM learning ineffective in understanding Mathematics. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -3.729, p = 0.0002, with a very low effect size (*r* = 0.07).

In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that only minority 20.6% of SIM teachers found SIM learning effective in understanding Mathematics. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -15.253, p = 0.0000, with a moderate effect size (*r* = 0.59).

15. **Effectiveness of SIM learning in understanding Dzongkha:** The 67.1% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in understanding Dzongkha subject in comparison to classroom learning. Looking at students' rating of SIM learning in understanding Dzongkha by age group, key stage, and school type, our data results show that all age groups and all key stages rated SIM learning "effective" in understanding Dzongkha. Similarly, the majority of school types rated SIM learning "effective" in understanding Dzongkha except ECR. ECR rated it as "ineffective." However, only 45.5% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in understanding Dzongkha.

    Evidence on effectiveness of SIM learning in understanding Dzongkha: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority 67.1% of SIM students found SIM learning effective in understanding Dzongkha subject. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 18.800, p = 0.0000, with a low effect size (*r* = 0.37).

    In the SIM teacher population, there is statistically significant evidence (p = 0.0080) that only minority 45.5% of SIM teachers, both female teachers and male teachers, found SIM learning effective in understanding Dzongkha. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -2.653, p = 0.0080, with a very low effect size (*r* = 0.10).

Advantages and Disadvantages of SIM Learning

16. **Advantages of SIM learning:** The SIM students found "Learning on your own pace" (62%) as the main advantage of SIM learning, followed by "Self-learning is fun" (57%) and "Ability to stay at home" (48%). Similarly, the SIM teachers found "Learning on your own pace" (79%) as the main advantage of SIM learning, followed by "Ability to stay at home" (47%) and "Self-learning is fun" (43%).

    Evidence on SIM students' perception on advantages of SIM learning: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority of SIM students found "Learning at your own pace" as the main advantage of SIM learning, followed by "Self-learning is fun". In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of advantages of SIM learning, $\chi^2(4, N = 2648) = 3604.269$, p =0.0000, with a large effect size ($\eta^2 = 0.34$). A pairwise post-hoc Cochran test was also significant for "Learning at your own pace" vs. "Self-learning is fun" (p = .0001) but the difference (effect size) between them is very small (η^2 = 0.01).

In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority of SIM teachers found "Learning at your own pace" as the main advantage of SIM learning, followed by "Ability to stay at home". In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of advantages of SIM learning, $\chi^2(4, N = 667) = 1073.172$, p =0.0000, with a large effect size (η^2 = 0.40). A pairwise post-hoc Cochran test was also significant for "Learning at your own pace" vs. "Ability to stay at home" (p = .0000) with a moderate difference (η^2 = 0.24).

17. **Disadvantages of SIM learning:** The SIM students found "Self-learning is difficult" (71%) as the main disadvantage of SIM learning, followed by "Household works at home" (49%) and "No self-discipline" (34%). Similarly, the SIM teachers found "Self-learning is difficult" (80%) as the main disadvantage of SIM learning, followed by "Household works at home" (52%) and "No self-discipline" (42%).

Evidence on SIM students' perception of disadvantages of SIM learning: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the majority of SIM students found "Self-learning is difficult" as the main and only disadvantage of SIM learning. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of disadvantages of SIM learning, $\chi^2(4, N = 2648) = 3558.177$, p =0.0000, with a large effect size (η^2 = 0.34). A pairwise post-hoc Cochran test was also significant for "Self-learning is difficult" vs. "Household works at home" (p = .0000) with a moderate effect size (η^2 = 0.09). Also, an interesting finding is that against conventional belief, "Household works at home" was not statistically significant disadvantage for the majority of students (p = 0.889581) as well as it is not true that girls were more affected than boys by household works (p = 0.4740) during SIM learning.

In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the majority of SIM teachers found "Self-learning is difficult" as the main disadvantage of SIM learning. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of disadvantages of SIM learning, $\chi^2(4, N = 667) = 1164.234$, p =0.0000, with a large effect size (η^2 = 0.44). A pairwise post-hoc Cochran test was also significant for "Self-learning is difficult" vs. "Household works at home" (p = .0000) with a moderate effect size (η^2 = 0.13).

Effect of Household Chores on SIM Learning

18. **Effect of Household Chores on SIM Learning: Is "Household works at home" a statistically significant disadvantage for the majority of the SIM students?**

One-sided binomial test indicated that the percentage of SIM students who selected "Household works at home" as a disadvantage (N_{hw} = 1293, 49%), was not statistically significantly different from the population hypothesized value of 50%, p = 0.889581 (which is much greater than alpha = 0.05). Therefore, there is no sufficient evidence that "Household works at home" affected the majority of SIM students during SIM learning.

Similarly, one-sided binomial test indicated that the percentage of SIM teachers who selected "Household works at home" as a disadvantage (N_{hw} = 346, 52%), was not statistically significantly different from the population hypothesized value of 50%, p = 0.176375 (which greater than alpha = 0.05). Therefore, there is no sufficient evidence that "Household works at home" affected the majority of SIM students during SIM learning even in the perception of SIM teachers.

19. **Gender Difference in Effect of Household Chores in SIM Learning: Is there gender difference in "Household works at home" for the SIM students?**

    Since our SIM survey sample is large enough (N=2648) to assume normal distribution, we applied two-sample test of proportions to test whether "Household works at home" affected girls more than boys during SIM learning in times of COVID-19 pandemic. We found that there is no statistically significant evidence that girls were affected more than boys by "Household works at home" during the SIM learning, z = 0.0651, p = 0.4740 (which is greater than alpha = 0.05). Therefore, "Household works at home" was not statistically significant disadvantage for the majority of students, both boys and girls, during SIM learning.

    Similarly, since our SIM survey sample is large enough (N=667) to assume normal distribution, we applied two-sample test of proportions to test whether "Household works at home" affected girls more than boys during SIM learning in times of COVID-19 pandemic. We found that there is no statistically significant evidence that girls were affected more than boys by "Household works at home" during the SIM learning, z = 0.0785, p = 0.4687 (which is greater than alpha = 0.05). Therefore, "Household works at home" was not statistically significant disadvantage for the in the perception of SIM teachers, both female teachers and male teachers, during SIM learning.

Help Sought for SIM Learning

20. **SIM students' help sought for SIM learning:** The 90.1% of SIM students said they sought help from someone to understand SIM lessons. The SIM students mainly sought help from teachers (44%) and siblings (44%), followed by student friends (39%) and parents (22%). Against a popular belief that SIM students would seek help from NFE instructors in the rural areas, only about 1% of the SIM students actually sought help from NFE instructors. About 10% of SIM students did not seek help from anyone. Similarly, the 94.6% of SIM teachers said they gave help to someone to understand SIM lessons. Likewise, the 99.2% of the SIM principal respondents said that their schools extended support to the SIM students. The 91.9% of the SIM principal respondents also said that their students or students' parents sought help regarding SIM. The 93.9% of the SIM parent respondents said that their schools offered help to their children. Also, the 92.0% of the SIM parent respondents said that their children sought help to understand SIM lessons.

Evidence on SIM students' help sought for SIM lessons: In the SIM student population, at least 89% of SIM students sought help for SIM lessons as there is statistically significant evidence (p = 0.035444) that the percentage of SIM students who sought help for SIM lessons is greater than population hypothesized value of 89%. In other words, a binomial test indicated that the percentage of SIM students who sought help for SIM lessons (N_{help} = 2386, 90.1%) was statistically significantly greater than the population hypothesized value of 89%, $p = 0.035444$ (which is less than significance level alpha = 0.05). Also, Cochran's Q test indicated that there are differences between the proportions among the five options of help for SIM lessons, $\chi^2(4, N = 2648) = 1670.831$, p =0.0000, with a large effect size ($\eta^2 = 0.16$). An exact pairwise post-hoc Cochran's Q test was not statistically significant for "Teacher" vs. "Sibling", $\chi^2(1, N = 2648) = 0.0191571$, $p = 0.9118$ (which is much greater than alpha = 0.05). Therefore, both teacher and sibling were equally number one helper for SIM lessons.

In the SIM teacher population, there is statistically significant evidence (p = 0.005874) that at least 92% of SIM teachers gave help for SIM lessons. In other words, a binomial test indicated that the percentage of SIM teachers who gave help for SIM lessons (N_{help} = 631, 94.6%) was statistically significantly greater than the population hypothesized value of 92%, $p = 0.005874$.

In the SIM principal population, there is statistically significant evidence (p = 0.013600) that at least 95% of SIM schools extended support to SIM students. A binomial test indicated that the percentage of SIM principals who believe their schools extended support to SIM students (N_{Yes} = 122, 99.2%), was statistically significantly greater than the population hypothesized value of 95%, $p = 0.013600$. Also in the SIM principal population, there is statistically significant evidence (p = 0.016869) that at least 85% of SIM students and parents sought help regarding SIM in the perception of principals. A binomial test indicated that the percentage of SIM principals who believe their students or students' parents sought help regarding SIM (N_{Yes} = 113, 91.9%), was statistically significantly greater than the population hypothesized value of 85%, $p = 0.016869$.

In the SIM parent population, there is statistically significant evidence (p = 0.028362) that at least 91% of SIM parents believe the schools offered help to their SIM children. A binomial test indicated that the percentage of SIM parents who believe their schools offered help to their SIM children (N_{Yes} = 351, 93.9%), was statistically significantly greater than the population hypothesized value of 91%, $p = 0.028362$. Also, in the SIM parent population, there is statistically significant evidence (p = 0.035098) that at least 89% of SIM parents believe their children sought help regarding SIM lessons. A binomial test indicated that the percentage of SIM parents who believe their children sought help to understand SIM lessons (N_{Yes} = 344, 92.0%), was statistically significantly greater than the population hypothesized value of 89%, $p = 0.035098$.

Comparison between SIM Learning and Classroom Learning

21. **Effectiveness of SIM learning vs Classroom Learning in increasing knowledge:** The 62.7% (SIM) vs 87.8% (Classroom) majority of the SIM student respondents rated

"effective" or "extremely effective" in increasing their knowledge. However, the 40.9% (SIM) vs 79.8% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in increasing knowledge.

Evidence on SIM students' perception of SIM learning vs Classroom Learning in increasing knowledge: In the SIM student population, there is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in increasing knowledge. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of increasing knowledge, Z = -29.089, p = 0.0000, with a moderate effect size or difference (*r* = 0.57).

In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in increasing knowledge. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of increasing knowledge, Z = -16.737, p = 0.0000, with a strong effect size or difference (*r* = 0.65).

22. **Effectiveness of SIM learning vs Classroom Learning in increasing skills:** The 56.9% (SIM) vs 85.7% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in increasing their skills. However, only 38.4% (SIM) vs 78.4% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in increasing skills.

    Evidence on SIM students' perception of SIM learning vs Classroom Learning in increasing skills: There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in increasing skills. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of increasing skills, Z = -26.939, p = 0.0000, with a moderate effect size or difference (*r* = 0.52).

    In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in increasing skills. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of increasing skills, Z = -16.489, p = 0.0000, with a strong effect size or strong difference (*r* = 0.64).

23. **Effectiveness of SIM learning vs Classroom Learning in imparting values:** The 54.6% (SIM) vs 85.1% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in imparting values. However, only 29.0% (SIM) vs 79.9% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in imparting values.

    Evidence on SIM students' perception of SIM learning vs Classroom Learning in imparting values: There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in imparting values.

In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of imparting values, $Z = -28.397$, $p = 0.0000$, with a moderate effect size or difference ($r = 0.55$).

In the SIM teacher population, there is statistically significant evidence ($p = 0.0000$) that the SIM teachers found classroom learning more effective than SIM learning in imparting values. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of imparting values, $Z = -17.976$, $p = 0.0000$, with a strong effect size or strong difference ($r = 0.70$).

24. **Effectiveness of SIM learning vs Classroom Learning in improving attitudes:** The 52.4% (SIM) vs 84.2% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in improving attitudes. However, only 23.1% (SIM) vs 77.9% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in improving attitudes.

    Evidence on SIM students' perception of SIM Learning vs Classroom Learning in improving attitudes: There is statistically significant evidence ($p = 0.0000$) that the SIM students found classroom learning more effective than SIM learning in improving attitudes. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of improving attitudes, $Z = -28.105$, $p = 0.0000$, with a moderate effect size or difference ($r = 0.55$).

    In the SIM teacher population, there is statistically significant evidence ($p = 0.0000$) that the SIM teachers found classroom learning more effective than SIM learning in improving attitudes. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of improving attitudes, $Z = -19.100$, $p = 0.0000$, with a strong effect size or strong difference ($r = 0.74$).

25. **Effectiveness of SIM learning vs Classroom Learning in understanding English:** The 56.6% (SIM) vs 86.7% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in understanding English. However, only 34.3% (SIM) vs 81.7% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in understanding English.

    Evidence on SIM students' perception of SIM Learning vs Classroom Learning in understanding English: There is statistically significant evidence ($p = 0.0000$) that the SIM students found classroom learning more effective than SIM learning in understanding English. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of understanding English, $Z = -28.962$, $p = 0.0000$, with a moderate effect size or difference ($r = 0.56$).

    In the SIM teacher population, there is statistically significant evidence ($p = 0.0000$) that the SIM teachers found classroom learning more effective than SIM learning in understanding English. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of understanding English, Z = -18.128, p = 0.0000, with a strong effect size or strong difference (r = 0.70).

26. **Effectiveness of SIM learning vs Classroom Learning in understanding Maths:** Only 47.9% (SIM) vs 81.4% (Classroom) of the SIM student respondents rated "effective" or "extremely effective" in understanding Mathematics. Similarly, only the 20.6% (SIM) vs 78.1% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in understanding Mathematics.

    Evidence on SIM students' perception of SIM Learning vs Classroom Learning in understanding Mathematics: There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in understanding Mathematics. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of understanding Mathematics, Z = -31.320, p = 0.0000, with a strong effect size or difference (r = 0.61).

    In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in understanding Mathematics. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of understanding Mathematics, Z = -19.116, p = 0.0000, with a strong effect size or difference (r = 0.74).

27. **Effectiveness of SIM learning vs Classroom Learning in understanding Dzongkha:** The 67.1% (SIM) vs 85.9% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in understanding Dzongkha. However, only 45.5% (SIM) vs 82.3% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in understanding Dzongkha.

    Evidence on SIM students' perception of SIM Learning vs Classroom Learning in understanding Dzongkha: There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in understanding Dzongkha. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of understanding Dzongkha, Z = -26.437, p = 0.0000, with a moderate effect size or difference (r = 0.51).

    In the SIM teacher population, there is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in understanding Dzongkha. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of understanding Dzongkha, Z = -16.950, p = 0.0000, with a strong effect size or strong difference (r = 0.66).

Implementation Effectiveness of SIM

28. **Perception on implementation of SIM:** The 78.1% of the SIM principal respondents rated that the SIM programme implementation was "effective" or "very effective." Similarly, the 93.1% of the SIM DEO respondents rated that the SIM programme implementation was "effective" or "very effective." The 86.8% of the SIM LG leader respondents rated that the SIM programme implementation was "effective" or "very effective." The 79.1% of the SIM parent respondents rated that the SIM programme implementation was "effective" or "very effective."

    Evidence on perception on implementation effectiveness of SIM: In the SIM principal population, there is statistically significant evidence ($p = 0.0000$) that 78.1% of SIM principals believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 6.594$, $p = 0.0000$, with a moderate effect size ($r = 0.59$).

    In the SIM DEO population, there is statistically significant evidence ($p = 0.0000$) that 93.1% of SIM DEOs believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 4.450$, $p = 0.0000$, with a very strong effect size ($r = 0.83$).

    In the SIM LG leader population, there is statistically significant evidence ($p = 0.0000$) that 86.8% of SIM LG leaders believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 6.314$, $p = 0.0000$, with a strong effect size ($r = 0.72$).

    In the SIM parent population, there is statistically significant evidence ($p = 0.0000$) that 79.1% of SIM parents believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 11.637$, $p = 0.0000$, with a moderate effect size ($r = 0.60$).

29. **Perception on delivery of SIM:** The 76.4% of the SIM principal respondents said that the Dzongkhag Education Office delivered the SIMs. Similarly, the 89.7% of the SIM DEO respondents said that the Dzongkhag Education Office delivered the SIMs. The 77.6% of the SIM LG respondents said that their gewog office provided support in delivering the SIMs. The 95.5% of the SIM parent respondents said that their children received SIM.

    Evidence on delivery of SIM: In the SIM principal population, there is statistically significant evidence ($p = 0.019772$) that at least 67.5% of SIM principals believe the Dzongkhag Education Office delivered the SIMs. A binomial test indicated that the percentage of SIM principals who believe the Dzongkhag Education Office delivered the SIMs ($N_{Yes} = 94$, 76.4%), was statistically significantly greater than the population hypothesized value of 67.5%, $p = 0.019772$.

In the SIM DEO population, there is statistically significant evidence (p = 0.0035460) that at least 74% of SIM DEOs believe the Dzongkhag Education Office delivered the SIMs. A binomial test indicated that the percentage of SIM DEOs who believe the Dzongkhag Education Office delivered the SIMs (N_{Yes} = 26, 89.7%), was statistically significantly greater than the population hypothesized value of 74%, p = 0.035460.

In the SIM LG leader population, there is statistically significant evidence (p = 0.029282) that at least 67% of SIM LG leaders believe their offices delivered the SIMs. A binomial test indicated that the percentage of SIM LG leaders who believe their offices supported in delivering the SIMs (N_{Yes} = 59, 77.6%), was statistically significantly greater than the population hypothesized value of 67%, p = 0.029282.

In the SIM parent population, there is statistically significant evidence (p = 0.033387) that at least 93% of SIM parents believe their children received SIM. A binomial test indicated that the percentage of SIM parents who believe their children received SIM (N_{Yes} = 357, 95.5%), was statistically significantly greater than the population hypothesized value of 93%, p = 0.033387.

30. **Perception on whether SIM reached the identified students:** The 93.5% of the SIM principal respondents said that the SIM has reached the identified students. Similarly, the 96.6% of the SIM DEO respondents said that the SIM has reached the identified students. The 97.4% of the SIM LG respondents said that the SIM has reached the identified students.

    Evidence on whether SIM reached the identified students: In the SIM principal population, there is statistically significant evidence (p = 0.023463) that at least 87.5% of SIM principals believe SIM has reached the identified students. A binomial test indicated that the percentage of SIM principals who believe the SIM has reached the identified students (N_{Yes} = 115, 93.5%), was statistically significantly greater than the population hypothesized value of 87.5%, p = 0.023463.

    In the SIM DEO population, there is statistically significant evidence (p = 0.041553) that at least 84% of SIM DEOs believe SIM has reached the identified students. A binomial test indicated that the percentage of SIM DEOs who believe the SIM has reached the identified students (N_{Yes} = 28, 96.6%), was statistically significantly greater than the population hypothesized value of 84%, p = 0.041553.

    In the SIM LG leader population, there is statistically significant evidence (p = 0.028065) that at least 91% of SIM LG leaders believe SIM has reached the identified students. A binomial test indicated that the percentage of SIM LG leaders who believe the SIM has reached the identified students (N_{Yes} = 74, 97.4%), was statistically significantly greater than the population hypothesized value of 91%, p = 0.028065.

31. **Perception on whether SIM reached other needy students:** The 87.0% of the SIM principal respondents said that the SIM has reached other needy students. Similarly, the

96.6% of the SIM DEO respondents said that the SIM has reached other needy students. The 88.2% of the SIM LG respondents said that the SIM has reached other needy students.

Evidence on whether SIM reached other needy students: In the SIM principal population, There is statistically significant evidence ($p = 0.021581$) that at least 79.5% of SIM principals believe SIM has reached other needy students. A binomial test indicated that the percentage of SIM principals who believe the SIM has reached other needy students ($N_{Yes} = 107$, 87.0%), was statistically significantly greater than the population hypothesized value of 79.5%, $p = 0.021581$.

In the SIM DEO population, there is statistically significant evidence ($p = 0.041553$) that at least 84% of SIM DEOs believe SIM has reached other needy students. A binomial test indicated that the percentage of SIM DEOs who believe the SIM has reached other needy students ($N_{Yes} = 28$, 96.6%), was statistically significantly greater than the population hypothesized value of 84%, $p = 0.041553$.

In the SIM LG leader population, there is statistically significant evidence ($p = 0.028670$) that at least 79% of SIM LG leaders believe SIM has reached other needy students. A binomial test indicated that the percentage of SIM LG leaders who believe the SIM has reached other needy students ($N_{Yes} = 67$, 88.2%), was statistically significantly greater than the population hypothesized value of 79%, $p = 0.028670$.

Introduction

On 6th March 2020, the Royal Government of Bhutan confirmed first case of COVID-19 in Bhutan, which resulted in the disruptions of face-to-face classroom learning in schools. All schools were obliged to adapt to Education in Emergency (EiE) curriculum from classes PP-XII. In this regard, the Ministry of Education (MoE) developed the Self-Instructional Materials (SIM) Programme with the theme "**Reaching the Unreached"** primarily to facilitate education of the students living in remote places with either limited or no access to *BBS* and *Internet* for e-learning lessons. After one year of SIM learning, the MoE decided to do a nationwide assessment study of SIM programme through perception surveys of SIM students, teachers, principals, district education officers (DEOs), parents and community leaders. Primarily, the quantitative survey method was used as main technique for data collection. In total, as shown in ***Table 1***, data were collected from samples of 2648 SIM students, 667 SIM teachers, 123 SIM principals, 29 chief DEOs and deputy DEOs, 374 SIM parents and 76 SIM local government leaders.

Table 1: SIM Data Collection Nationwide

Sl.	Target Population	Sample Size
1.	SIM Students	2648
2.	SIM Teachers	667
3.	SIM Principals	123
4.	SIM DEOs	29
5.	SIM Community Leaders	76
6.	SIM Parents	374

The main target population for the SIM assessment study was SIM students for which we were able to get large and nationwide representative sample size of 2648 SIM students, for external validity and generalizability of our evidence findings. To support main target population of SIM students' perception on SIM programme, perceptions of SIM teachers, SIM principals, DEOs, SIM parents and SIM local government leaders were collected too. SIM data were collected by more than 120 trained SIM teacher enumerators and data were collected from more than 80 schools in all 20 Dzongkhags nationwide, consisting of all types of school such as higher secondary schools (HSS), middle secondary schools (MSS), lower secondary schools (LSS), primary schools (PS) and extended classrooms (ECR). For data honesty, integrity and quality, all respondents were informed about the objectives of the study and agreed to voluntarily participate. Moreover, data were collected anonymously with no individual identifying information collected. The study questionnaires were approved by the Ministry of Education. For SIM data sampling technique, stratified random sampling strategy was used for gender representation as well as for representation across all classes and all key stages of SIM materials which have five key stages. Before the main data collection, questionnaires were pre-tested for any technical problems as well as for any ethical sensitivity. Pre-testing were done on 210 SIM students, 107 SIM teachers, 44 SIM parents and 15 local government leaders. Based on the feedbacks from pre-testing, technical adjustments were made as well as suggestions were incorporated to reflect ground reality. Similarly, once data were collected, data cleaning and data coding works were carried out carefully including spotting data outliers before data were analyzed using statistical software STATA 17.0 for evidence findings.

PART I: SIM STUDENTS

Demographic Characteristics of SIM Student Respondents

The age characteristics of the SIM student respondents are summarized in ***Table 2***. The age of the SIM student respondents ranged from 6 to 24 years (M = 13.38, SD = 3.50).

Table 2: Results of age characteristics of SIM student respondents

Variable	Obs	Mean	Std. dev.	Min	Max
age	2,648	13.37689	3.501301	6	24

Similarly, among the 2648 SIM student respondents, 1210 (45.7%) were males and 1438 (54.3%) were females as shown in ***Figure 1***.

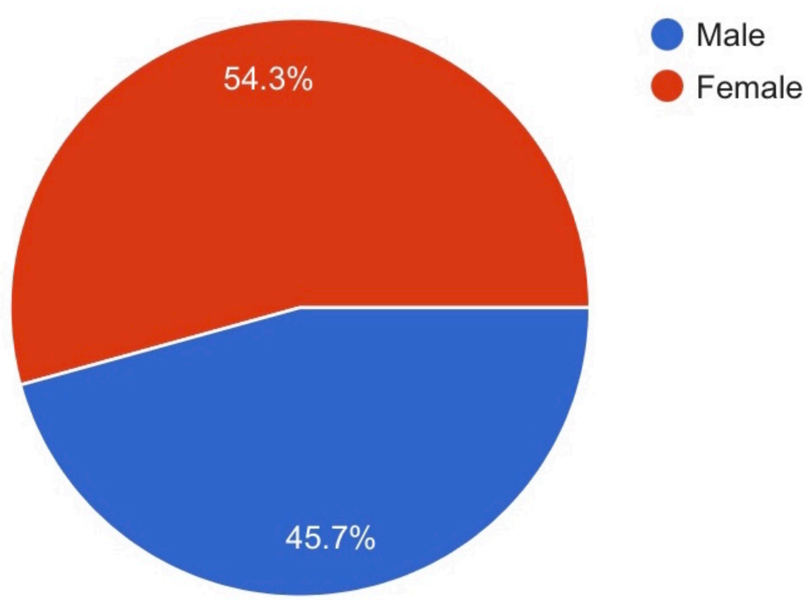


Figure 1: *Gender of SIM student respondents*

Likewise, among the 2648 SIM student respondents, we got data representation from all classes from Class I to Class XII as shown in ***Figure 2,*** with maximum from class VI (12.0%), closely followed by class X (11.4%), class VII (10.2%), class IX(10.2%), class V (9.7%), class IV (7.9%), class XII (7.4%), class III (6.7%), class XI (6.6%), class VIII (6.4%), class II (6.1%) and with minimum from class I (5.4%). Class PP students were not surveyed because they did not exist last year (2020) when SIM programme was implemented.

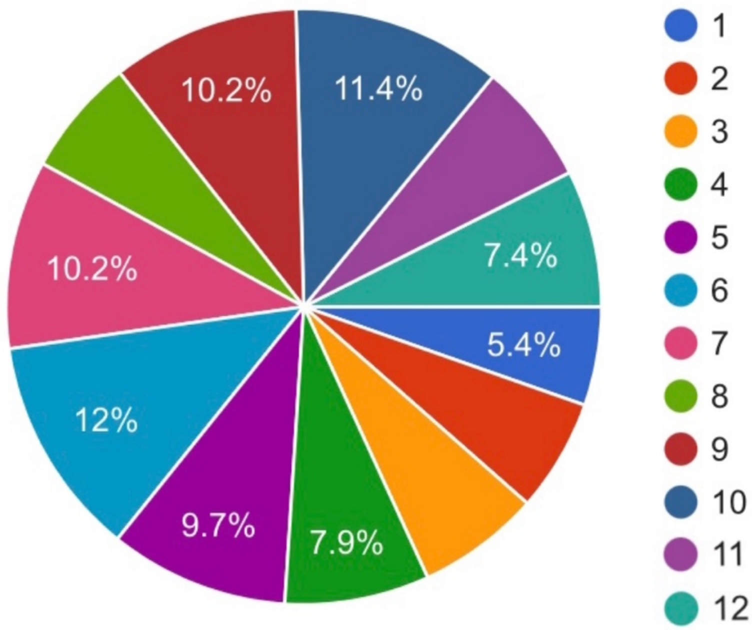


Figure 2*: Classes of SIM student respondents*

Among the 2648 SIM student respondents, we got data representation from all types of schools as shown in ***Figure 3,*** with maximum from HSS (36.0%), followed by PS (30.7%), MSS (24.1%), LSS (6.7%), and with minimum from ECR (2.6%).

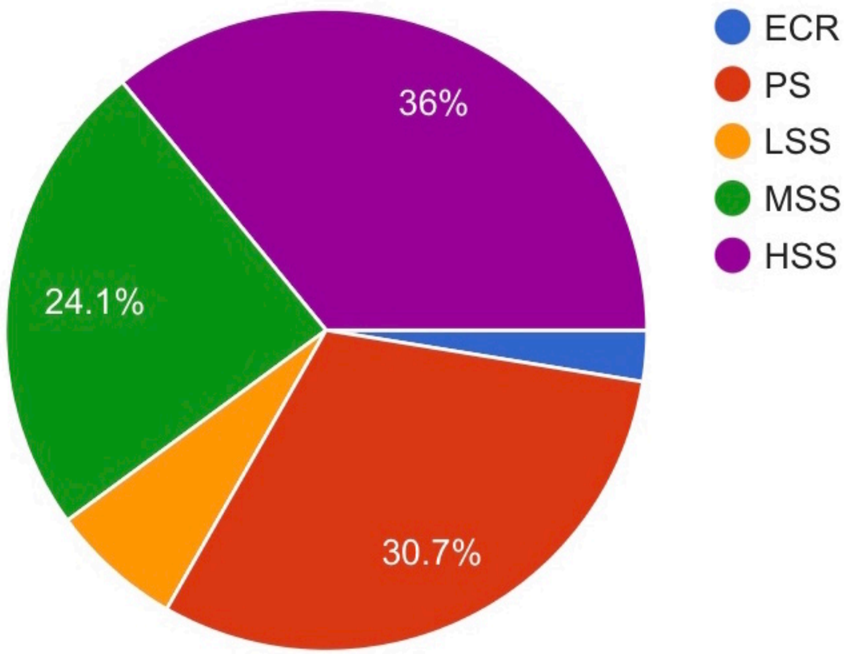


Figure 3*: School types of SIM student respondents*

Effectiveness of SIM Programme

Analyzing Students' Satisfaction Level of SIM

The Ministry of Education was interested to know satisfaction level of SIM programme, especially SIM students' satisfaction level in particular, during COVID-19 pandemic. To investigate this, ***Figure 4***, which is visualization of survey data, shows the results of satisfaction opinion from the SIM survey.

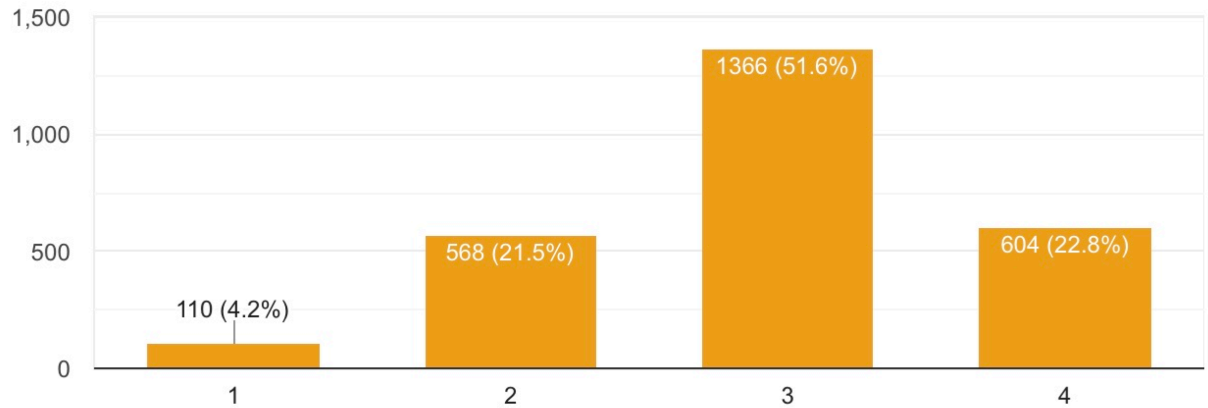


Figure 4*: Results of "Rate how satisfied are you with the current SIM" where 1 = Extremely dissatisfied, 2 = Dissatisfied, 3 = Satisfied, and 4 = Extremely satisfied*

As can be seen in ***Figure 4*** the 74.4% of the SIM student respondents rated the SIM programme "satisfied" or "extremely satisfied."

Descriptive Analysis – Measure of Central Tendency

Table 3: Results of the SIM students' satisfaction level rating frequency distribution

```
. tabulate q27

        q27 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        110        4.15        4.15
          2 |        568       21.45       25.60
          3 |      1,366       51.59       77.19
          4 |        604       22.81      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency ***Table 3*** above, it shows that mode choice is 3, which is "satisfied." The total SIM student respondents of 74.4% chose "satisfied" or "extremely satisfied."

Table 4: SIM students' satisfaction level rating frequency distribution, by age group

```
. tabulate age_group q27
```

Age_Group	q27 1	2	3	4	Total
(10-14)	30	214	620	322	1,186
(15-19)	66	260	506	151	983
(20-24)	5	22	38	5	70
(5-9)	9	72	202	126	409
Total	110	568	1,366	604	2,648

Looking at students' satisfaction level of SIM survey data by age group, it shows that consistently in all age groups, the mode or most choice selected is 3, which is "satisfied."

Table 5: SIM students' satisfaction level rating frequency distribution, by key stage

```
. tabulate key_stage q27
```

Key_Stage	q27 1	2	3	4	Total
Key Stage I	9	84	237	151	481
Key Stage II	6	129	408	240	783
Key Stage III	32	102	233	74	441
Key Stage IV	35	168	287	81	571
Key Stage V	28	85	201	58	372
Total	110	568	1,366	604	2,648

Similarly, looking at students' satisfaction level of SIM survey data by key stage, it shows that consistently in all key stages, the mode or most choice selected is 3, which is "satisfied."

Table 6: SIM students' satisfaction level rating frequency distribution, by school type

```
. tabulate school q27
```

School	q27 1	2	3	4	Total
ECR	1	2	38	27	68
HSS	65	256	478	153	952
LSS	3	44	93	36	176
MSS	32	146	349	112	639
PS	9	120	408	276	813
Total	110	568	1,366	604	2,648

Likewise, looking at students' satisfaction level of SIM survey data by school type, it shows that consistently in all school types, the mode or most choice selected is 3, which is "satisfied."

Table 7: Result of the SIM students' satisfaction level rating median calculation

```
. tabstat q27, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q27 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "satisfied." This means at least 50% of the SIM student respondents are in the "satisfied" or "extremely satisfied" category looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 8: Result of the SIM students' measure of consensus on satisfaction level

```
. cns q27 , min(1) max(4)

Consensus Measure for q27
Cns(X) = .66077408
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the satisfaction level of SIM students, it is 0.6608.

Inferential Analysis – Statistical Significance Testing

Table 9: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q27 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1970     2705671     1753638
    Negative |      678      801605     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.583e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.390e+09

H0: q27 = 2.5
         z = 25.537
Prob > |z| = 0.0000
```

We have seen that the 74.4% of SIM students surveyed think that SIM programme was satisfactory. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "dissatisfied" and 3 = "satisfied."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 25.537, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 25.537 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 25.537/SQRT(2648) = 0.50. This, according to Bartz (1999) is moderate effect size.

Gender difference in satisfaction level of SIM learning

Table 10: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q27, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1921033     1904631
        Male |     1210     1586243     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -61104843
                     ----------
Adjusted variance     3.230e+08

H0: q27(gender==Female) = q27(gender==Male)
         z =  0.913
Prob > |z| = 0.3614
```

There is no evidence for statistically significant difference between satisfaction level of SIM learning between female students and male students (p-value = 0.3614 > alpha = 0.05), which means both girls and boys are equally satisfied with SIM learning.

Evidence on SIM Students' Satisfaction Level

There is statistically significant evidence (p = 0.0000) that the majority 74.4% of SIM students, both female students and male students, are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 25.537, p = 0.0000, with a moderate effect size (r = 0.50).

Analyzing Students' Acceptance Level of SIM

The Ministry of Education was interested to know acceptance level of SIM programme, especially SIM students' acceptance level in particular, during COVID-19 pandemic. To investigate this, ***Figure 5*** shows the results of SIM acceptance opinion from the SIM survey.

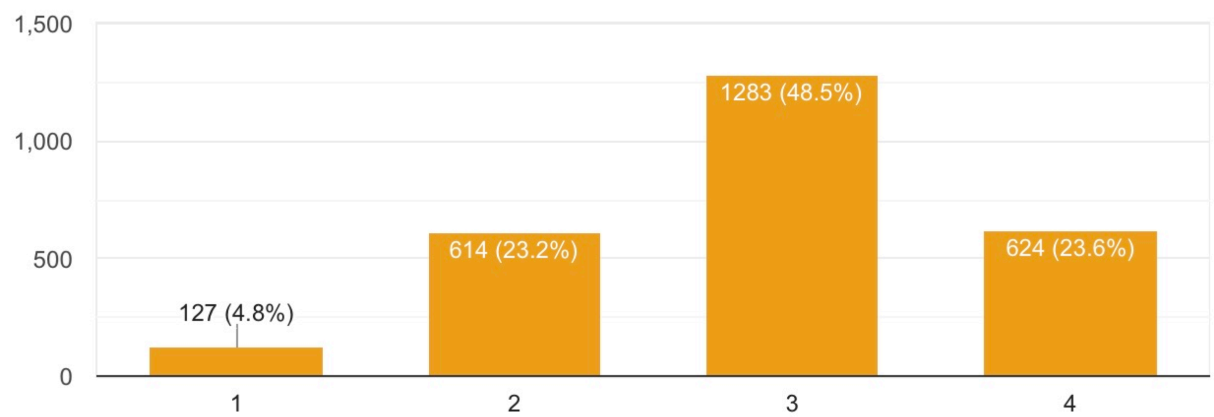


Figure 5*: Results of "Rate how much did you enjoy SIM learning during the pandemic" where 1 = Extremely unenjoyable, 2 = Unenjoyable, 3 = Enjoyable, and 4 = Extremely enjoyable*

As can be seen in ***Figure 5*** the 72.1% of the SIM student respondents rated the SIM learning "enjoyable" or "extremely enjoyable."

Descriptive Analysis – Measure of Central Tendency

Table 11: Results of the SIM students' acceptance level rating frequency distribution

```
. tabulate q21

        q21 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        127        4.80        4.80
          2 |        614       23.19       27.98
          3 |      1,283       48.45       76.44
          4 |        624       23.56      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "enjoyable." The total SIM student respondents of 72.1% chose "enjoyable" or "extremely enjoyable."

Table 12: SIM students' acceptance level rating frequency distribution, by age group

```
. tabulate age_group q21

           |                    q21
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (10-14) |        38        233        608        307 |     1,186
   (15-19) |        67        272        474        170 |       983
   (20-24) |         6         21         28         15 |        70
     (5-9) |        16         88        173        132 |       409
-----------+--------------------------------------------+----------
     Total |       127        614      1,283        624 |     2,648
```

Looking at students' acceptance level of SIM survey data by age group, it shows that consistently in all age groups, the mode or most choice selected is 3, which is "enjoyable."

Table 13: SIM students' acceptance level rating frequency distribution, by key stage

```
. tabulate key_stage q21

               |                    q21
     Key_Stage |         1          2          3          4 |     Total
---------------+--------------------------------------------+----------
   Key Stage I |        17         96        207        161 |       481
  Key Stage II |        17        128        418        220 |       783
 Key Stage III |        29        132        208         72 |       441
  Key Stage IV |        41        162        278         90 |       571
   Key Stage V |        23         96        172         81 |       372
---------------+--------------------------------------------+----------
         Total |       127        614      1,283        624 |     2,648
```

Similarly, looking at students' acceptance level of SIM survey data by key stage, it shows that consistently in all key stages, the mode or most choice selected is 3, which is "enjoyable."

Table 14: SIM students' acceptance level rating frequency distribution, by school type

```
. tabulate school q21

           |                        q21
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         2          9         36         21 |        68
       HSS |        66        276        434        176 |       952
       LSS |         4         41         92         39 |       176
       MSS |        39        165        330        105 |       639
        PS |        16        123        391        283 |       813
-----------+--------------------------------------------+----------
     Total |       127        614      1,283        624 |     2,648
```

Table 15: Result of the SIM students' acceptance level rating median calculation

Likewise, looking at students' acceptance level of SIM survey data by school type, it shows that consistently in all school types, the mode or most choice selected is 3, which is "enjoyable."

```
. tabstat q21, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q21 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "enjoyable." This means at least 50% of the SIM student respondents are in the "enjoyable" or "extremely enjoyable" group looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 16: Result of the SIM students' measure of consensus on acceptance level

```
. cns q21 , min(1) max(4)

Consensus Measure for q21
Cns(X) = .63382518
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the acceptance level of SIM students, it is 0.6338.

Inferential Analysis – Statistical Significance Testing

Table 17: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q21 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1907     2635919     1753638
    Negative |      741      871357     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.510e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.397e+09

H0: q21 = 2.5
         z = 23.604
Prob > |z| = 0.0000
```

We have seen that the 72.1% of SIM students surveyed think that SIM programme was enjoyable or extremely enjoyable. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "unenjoyable" and 3 = "enjoyable."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 23.604, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 23.604 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 23.604/SQRT(2648) = 0.46. This, according to Bartz (1999) is moderate effect size.

Gender difference in acceptance level of SIM learning

Table 18: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q21, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438   1916896.5     1904631
        Male |     1210   1590379.5     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -53545937
                     ----------
Adjusted variance     3.306e+08

H0: q21(gender==Female) = q21(gender==Male)
         z =  0.675
Prob > |z| = 0.4999
```

There is no evidence for statistically significant difference between acceptance level of SIM learning between female students and male students (p-value = 0.4999 > alpha = 0.05), which means both girls and boys found SIM learning equally enjoyable.

Evidence on SIM Students' Acceptance Level

There is statistically significant evidence (p = 0.0000) that the majority 72.1% of SIM students, both girls and boys, found SIM learning enjoyable during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 23.604, p = 0.0000, with a moderate effect size (*r* = 0.46).

Effectiveness of SIM Materials

Analyzing Students' Perception on Overall Presentation of SIM Booklets

The Ministry of Education was interested to know how SIM students found overall presentation of the SIM booklets. To investigate this, ***Figure 6*** shows the results of SIM students' perception on overall presentation of the SIM booklets.

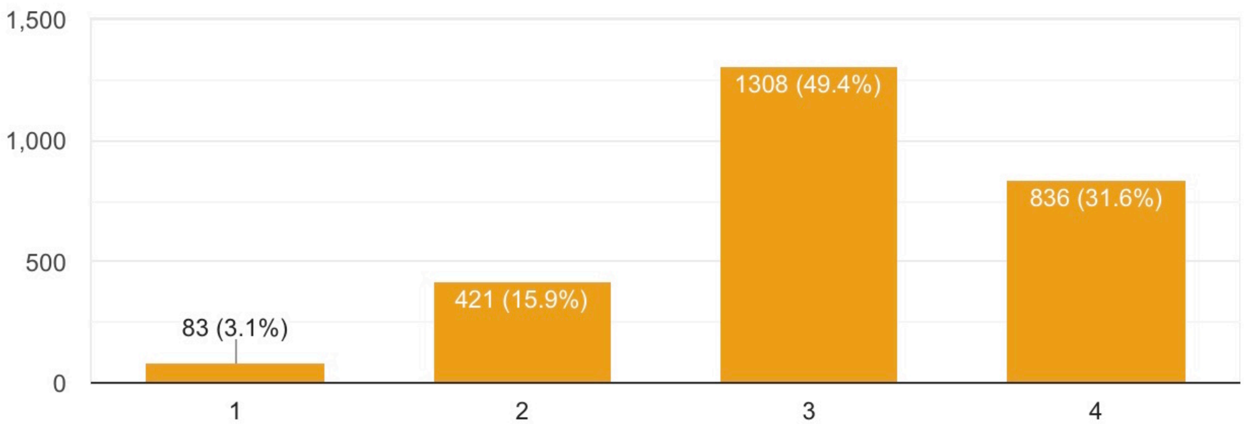


Figure 6*: Results of "Rate how did you find overall presentation of the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 6*** the 81.0% of the SIM student respondents rated the overall presentation of SIM materials "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 19: Results of the SIM students' rating of overall presentation of SIM frequency distribution

```
. tabulate q26

        q26 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         83        3.13        3.13
          2 |        421       15.90       19.03
          3 |      1,308       49.40       68.43
          4 |        836       31.57      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode is 3, which is "effective." The total SIM student respondents of 81.0% chose "effective" or "extremely effective."

Table 20: SIM students' rating of SIM overall presentation frequency distribution, by age group

```
. tabulate age_group q26

           |                    q26
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (10-14) |        31        153        593        409 |     1,186
   (15-19) |        44        208        471        260 |       983
   (20-24) |         2         16         36         16 |        70
     (5-9) |         6         44        208        151 |       409
-----------+--------------------------------------------+----------
     Total |        83        421      1,308        836 |     2,648
```

Looking at students' rating of SIM overall presentation by age group, it shows consistently that all age groups have mode 3, which is "effective."

Table 21: SIM students' rating of SIM overall presentation frequency distribution, by key stage

```
. tabulate key_stage q26

              |                    q26
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         6         53        239        183 |       481
 Key Stage II |        13         88        393        289 |       783
Key Stage III |        25         79        224        113 |       441
 Key Stage IV |        26        134        280        131 |       571
  Key Stage V |        13         67        172        120 |       372
--------------+--------------------------------------------+----------
        Total |        83        421      1,308        836 |     2,648
```

Similarly, looking at students' rating of SIM overall presentation by key stage, it shows consistently that all key stages have mode 3, which is "effective."

Table 22: SIM students' rating of SIM overall presentation frequency distribution, by school type

```
. tabulate school q26

           |                    q26
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          3         22         43 |        68
       HSS |        53        203        438        258 |       952
       LSS |         3         33        101         39 |       176
       MSS |        18         94        349        178 |       639
        PS |         9         88        398        318 |       813
-----------+--------------------------------------------+----------
     Total |        83        421      1,308        836 |     2,648
```

Likewise, looking at students' rating of SIM overall presentation by school type, it shows that majority of school types such as HSS, MSS, LSS, and PS rated SIM overall presentation as "effective" with mode of 3 while ECR rated SIM overall presentation as "extremely effective" with mode of 4.

Table 23: Result of the SIM students' rating of SIM overall presentation median calculation

```
. tabstat q26, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q26 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM overall presentation "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 24: Result of the SIM students' measure of consensus on SIM overall presentation rating

```
. cns q26 , min(1) max(4)

Consensus Measure for q26
Cns(X) = .65536028
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM overall presentation rating of SIM students, it is 0.6553.

Inferential Analysis – Statistical Significance Testing

Table 25: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q26 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     2144     2961424     1753638
    Negative |      504      545852     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.239e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.424e+09

H0: q26 = 2.5
         z = 32.003
Prob > |z| = 0.0000
```

We have seen that the 81.0% of SIM students surveyed think that SIM overall presentation was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 32.003, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 32.003 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 32.003/SQRT(2648) = 0.62. This, according to Bartz (1999) is strong effect size.

Gender difference in SIM students' rating of SIM overall presentation

Table 26: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q26, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1897156     1904631
        Male |     1210     1610120     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -59935036
                     ----------
Adjusted variance     3.242e+08

H0: q26(gender==Female) = q26(gender==Male)
         z = -0.415
Prob > |z| = 0.6780
```

There is no evidence for statistically significant difference in SIM overall presentation rating between female students and male students (p-value = 0.6780 > alpha = 0.05), which means both girls and boys found SIM overall presentation equally effective.

Evidence on SIM Students' Perception of SIM Overall Presentation

There is statistically significant evidence (p = 0.0000) that the majority 81.0% of SIM students, both girls and boys, found overall presentation of the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 32.003, p = 0.0000, with a strong effect size (r = 0.62).

Analyzing Students' Perception on Contents in SIM Booklets

The Ministry of Education was interested to know how SIM students found contents of the SIM booklets. To investigate this, ***Figure 7*** shows the results of SIM students' perception on contents of the SIM booklets.

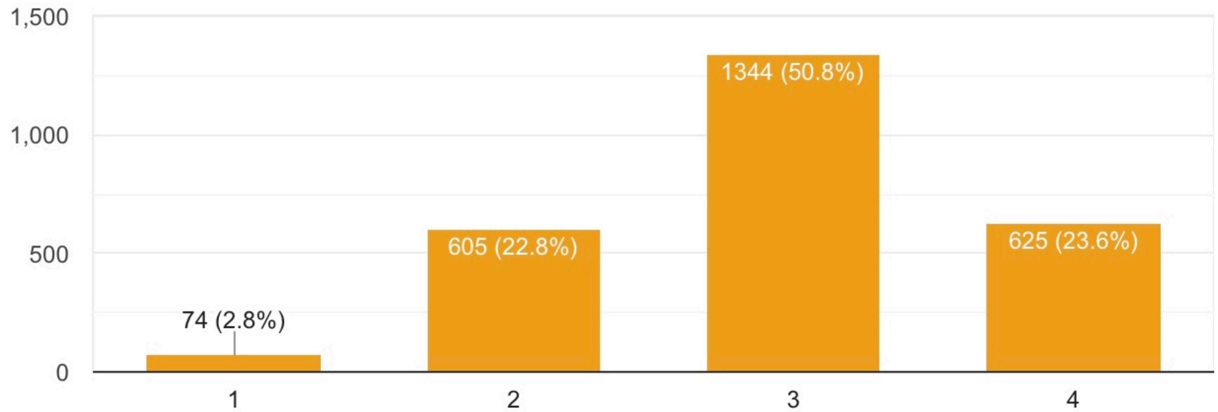


***Figure** 7: Results of "Rate how did you find contents of the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure** 7* the 74.4% of the SIM student respondents rated the SIM contents "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 27: Results of the SIM students' rating of SIM contents frequency distribution

```
. tabulate q23

        q23 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         74        2.79        2.79
          2 |        605       22.85       25.64
          3 |      1,344       50.76       76.40
          4 |        625       23.60      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 74.4% chose "effective" or "extremely effective."

Table 28: SIM students' rating of SIM contents frequency distribution, by age group

```
. tabulate age_group q23

           |                         q23
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (10-14) |        33        215        622        316 |     1,186
   (15-19) |        34        258        508        183 |       983
   (20-24) |         4         25         32          9 |        70
     (5-9) |         3        107        182        117 |       409
-----------+--------------------------------------------+----------
     Total |        74        605      1,344        625 |     2,648
```

Looking at students' rating of SIM contents by age group, it shows that in all age groups the mode choice selected is 3, which is "effective."

Table 29: SIM students' rating of SIM contents frequency distribution, by key stage

```
. tabulate key_stage q23

              |                         q23
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         3        118        217        143 |       481
 Key Stage II |         9        123        426        225 |       783
Key Stage III |        31        110        221         79 |       441
 Key Stage IV |        17        165        279        110 |       571
  Key Stage V |        14         89        201         68 |       372
--------------+--------------------------------------------+----------
        Total |        74        605      1,344        625 |     2,648
```

Similarly, looking at students' rating of SIM contents by key stage, it shows that consistently in all key stages, the mode is 3, which is "effective."

Table 30: SIM students' rating of SIM contents frequency distribution, by school type

```
. tabulate school q23

           |                         q23
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0         10         26         32 |        68
       HSS |        46        270        459        177 |       952
       LSS |         7         35        103         31 |       176
       MSS |        16        157        346        120 |       639
        PS |         5        133        410        265 |       813
-----------+--------------------------------------------+----------
     Total |        74        605      1,344        625 |     2,648
```

Likewise, looking at students' rating of SIM contents by school type, it shows that consistently in majority school types, the mode is 3, which is "effective" and in the case of ECR, the mode is 4, which is "extremely effective."

Table 31: Result of the SIM students' rating of SIM contents median calculation

```
. tabstat q23, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q23 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM contents "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 32: Result of the SIM students' measure of consensus on SIM contents rating

```
. cns q23 , min(1) max(4)

Consensus Measure for q23
Cns(X) = .67354071
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM contents rating of SIM students, it is 0.6735.

Inferential Analysis – Statistical Significance Testing

Table 33: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q23 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1969     2747275     1753638
    Negative |      679      760001     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.614e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.387e+09

H0: q23 = 2.5
         z = 26.682
Prob > |z| = 0.0000
```

We have seen that the 74.4% of SIM students surveyed think that SIM contents was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 26.682, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 26.682 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 26.682/SQRT(2648) = 0.52. This, according to Bartz (1999) is moderate effect size.

Gender difference in SIM students' rating of SIM contents

Table 34: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q23, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438   1912151.5     1904631
        Male |     1210   1595124.5     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -59861167
                     ----------
Adjusted variance     3.242e+08

H0: q23(gender==Female) = q23(gender==Male)
         z =  0.418
Prob > |z| = 0.6762
```

There is no evidence for statistically significant difference between SIM contents rating between female students and male students (p-value = 0.6762 > alpha = 0.05), which means both girls and boys found SIM contents equally effective.

Evidence on SIM Students' Perception of SIM Contents

There is statistically significant evidence (p = 0.0000) that the majority 74.4% of SIM students, both girls and boys, found contents of SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 26.682, p = 0.0000, with a moderate effect size (*r* = 0.52).

Analyzing Students' Perception on Instructions in SIM Booklets

The Ministry of Education was interested to know how SIM students found instructions incorporated in the SIM booklets. To investigate this, ***Figure 8*** shows the results of SIM students' perception on instructions in the SIM booklets.

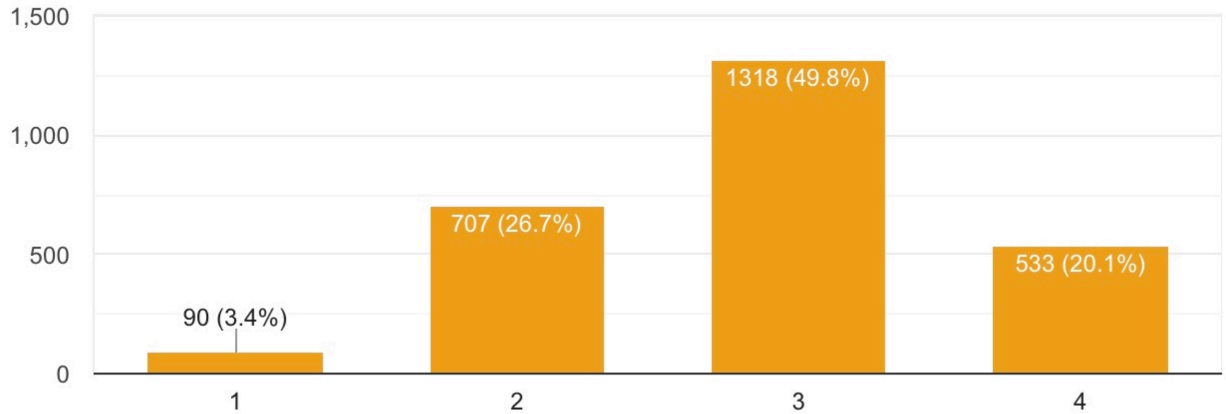


Figure 8*: Results of "Rate how did you find instructions in the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 8*** the 69.9% of the SIM student respondents rated the SIM instructions "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 35: Results of the SIM students' rating of SIM instructions frequency distribution

```
. tabulate q22

        q22 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         90        3.40        3.40
          2 |        707       26.70       30.10
          3 |      1,318       49.77       79.87
          4 |        533       20.13      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 69.9% chose "effective" or "extremely effective."

Table 36: SIM students' rating of SIM instructions frequency distribution, by age group

```
. tabulate age_group q22
```

Age_Group	q22 1	2	3	4	Total
(10-14)	32	265	618	271	1,186
(15-19)	43	291	485	164	983
(20-24)	6	26	24	14	70
(5-9)	9	125	191	84	409
Total	90	707	1,318	533	2,648

Looking at students' rating of SIM instructions by age group, it shows that in all age groups except age group 20-24, the mode or most choice selected is 3, which is "effective." The age group 20-24 has mode as 2, which is ineffective but difference between frequency of 2 (ineffective) and 3 (effective) is marginal 26 vs 24, which does not look significant.

Table 37: SIM students' rating of SIM instructions frequency distribution, by key stage

```
. tabulate key_stage q22
```

Key_Stage	q22 1	2	3	4	Total
Key Stage I	10	139	229	103	481
Key Stage II	13	171	399	200	783
Key Stage III	24	113	238	66	441
Key Stage IV	22	189	280	80	571
Key Stage V	21	95	172	84	372
Total	90	707	1,318	533	2,648

Similarly, looking at students' rating of SIM instructions by key stage, it shows that consistently in all key stages, the mode is 3, which is "effective."

Table 38: SIM students' rating of SIM instructions frequency distribution, by school type

```
. tabulate school q22

           |                      q22
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1         12         27         28 |        68
       HSS |        56        277        441        178 |       952
       LSS |         4         43        104         25 |       176
       MSS |        19        189        342         89 |       639
        PS |        10        186        404        213 |       813
-----------+--------------------------------------------+----------
     Total |        90        707      1,318        533 |     2,648
```

Likewise, looking at students' rating of SIM instructions by school type, it shows that consistently in all school types, the mode is 3, which is "effective."

Table 39: Result of the SIM students' rating of SIM instructions median calculation

```
. tabstat q22, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q22 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM instructions "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 40: Result of the SIM students' measure of consensus on SIM instructions rating

```
. cns q22 , min(1) max(4)

Consensus Measure for q22
Cns(X) = .65045756
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM instructions rating of SIM students, it is 0.6505.

Inferential Analysis – Statistical Significance Testing

Table 41: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q22 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1851     2580755     1753638
    Negative |      797      926521     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.780e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.370e+09

H0: q22 = 2.5
         z = 22.345
Prob > |z| = 0.0000
```

We have seen that the 69.9% of SIM students surveyed think that SIM instructions was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 22.345, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 22.345 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 22.345/SQRT(2648) = 0.43. This, according to Bartz (1999) is moderate effect size.

Gender difference in SIM students' rating of SIM instructions

Table 42: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q22, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1926147     1904631
        Male |     1210     1581129     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -57820726
                     ----------
Adjusted variance     3.263e+08

H0: q22(gender==Female) = q22(gender==Male)
         z =  1.191
Prob > |z| = 0.2336
```

There is no evidence for statistically significant difference between SIM instructions rating between female students and male students (p-value = 0.2336 > alpha = 0.05), which means both girls and boys found SIM instructions equally effective.

Evidence on SIM Students' Perception on SIM Instructions

There is statistically significant evidence (p = 0.0000) that the majority 69.9% of SIM students, both girls and boys, found instructions in SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 22.345, p = 0.0000, with a moderate effect size (*r* = 0.43).

Analyzing Students' Perception on Graphics in SIM Booklets

The Ministry of Education was interested to know how SIM students found graphics in the SIM booklets. To investigate this, ***Figure 9*** shows the results of SIM students' perception on graphics in the SIM booklets.

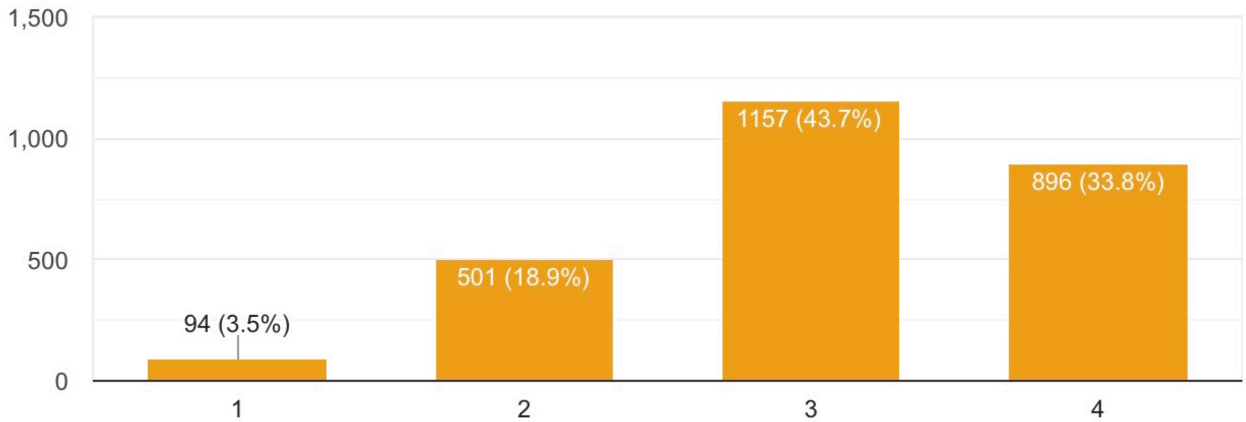


Figure 9*: Results of "Rate how did you find graphics in the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 9*** the 77.5% of the SIM student respondents rated the SIM graphics "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 43: Results of the SIM students' rating of SIM graphics frequency distribution

```
. tabulate q24

        q24 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         94        3.55        3.55
          2 |        501       18.92       22.47
          3 |      1,157       43.69       66.16
          4 |        896       33.84      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 77.5% chose "effective" or "extremely effective."

Table 44: SIM students' rating of SIM graphics frequency distribution, by age group

```
. tabulate age_group q24
```

Age_Group	q24: 1	2	3	4	Total
(10-14)	21	162	533	470	1,186
(15-19)	58	268	433	224	983
(20-24)	7	19	28	16	70
(5-9)	8	52	163	186	409
Total	94	501	1,157	896	2,648

Looking at students' rating of SIM graphics by age group, it shows that in majority age groups the mode is 3, which is "effective." Interestingly, the youngest age group of 5-9 year old rated SIM graphics "extremely effective" as they have mode of 4. It seems graphics in the SIM booklets were appreciated more by the younger children than the older children, although older children also rated them "effective."

Table 45: SIM students' rating of SIM graphics frequency distribution, by key stage

```
. tabulate key_stage q24
```

Key_Stage	q24: 1	2	3	4	Total
Key Stage I	9	60	194	218	481
Key Stage II	7	90	344	342	783
Key Stage III	22	84	210	125	441
Key Stage IV	32	162	253	124	571
Key Stage V	24	105	156	87	372
Total	94	501	1,157	896	2,648

Similarly, looking at students' rating of SIM graphics by key stage, it shows that majority of key stages have the mode as 3, which is "effective." Interestingly, consistent with how the youngest age group of 5-9 year old rated SIM graphics "extremely effective," the key stage I also rated SIM graphics "extremely effective" as they have mode of 4 . It seems graphics in the SIM booklets were appreciated more by the younger children than the older children or appreciated more by the lower classes than the higher classes, although all higher key stages also rated them "effective." This will have an important policy implication for the future material designs of SIM booklets that it's more effective to include more graphics for lower classes.

Table 46: SIM students' rating of SIM graphics frequency distribution, by school type

```
. tabulate school q24
```

School	q24 1	2	3	4	Total
ECR	1	8	22	37	68
HSS	59	256	388	249	952
LSS	3	39	89	45	176
MSS	23	111	330	175	639
PS	8	87	328	390	813
Total	94	501	1,157	896	2,648

Likewise, looking at students' rating of SIM graphics by school type, it shows that higher class level schools such as HSS, MSS and LSS rated SIM graphics as "effective" with mode of 3 while lower class level schools such as ECR and PS rated SIM graphics as "extremely effective" with mode of 4. This is consistent with how the youngest age group of 5-9 year old rated SIM graphics "extremely effective" and how the lowest key stage I rated SIM graphics "extremely effective." This is an important and consistent finding.

Table 47: Result of the SIM students' rating of SIM graphics median calculation

```
. tabstat q24, stat(count p50 min max)
```

Variable	N	p50	Min	Max
q24	2648	3	1	4

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM graphics "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 48: Result of the SIM students' measure of consensus on SIM graphics rating

```
. cns q24 , min(1) max(4)

Consensus Measure for q24
Cns(X) = .62215072
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM graphics rating of SIM students, it is 0.6222.

Inferential Analysis – Statistical Significance Testing

Table 49: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q24 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     2053   2889267.5     1753638
    Negative |      595    618008.5     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.152e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.433e+09

H0: q24 = 2.5
         z = 29.999
Prob > |z| = 0.0000
```

We have seen that the 77.5% of SIM students surveyed think that SIM graphics was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 29.999, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 29.999 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 29.999/SQRT(2648) = 0.58. This, according to Bartz (1999) is moderate effect size.

Gender difference in SIM students' rating of SIM graphics

Table 50: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q24, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1879925     1904631
        Male |     1210     1627351     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -49538829
                     ----------
Adjusted variance     3.346e+08

H0: q24(gender==Female) = q24(gender==Male)
         z = -1.351
Prob > |z| = 0.1768
```

There is no evidence for statistically significant difference between SIM graphics rating between female students and male students (p-value = 0.1768 > alpha = 0.05), which means both girls and boys found SIM graphics equally effective.

Evidence on SIM Students' Perception of SIM Graphics

There is statistically significant evidence (p = 0.0000) that the majority 77.5% of SIM students, both girls and boys, found graphics in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 29.999, p = 0.0000, with a moderate effect size (*r* = 0.58).

Analyzing Students' Perception on Activities in SIM Booklets

The Ministry of Education was interested to know how SIM students found activities in the SIM booklets. To investigate this, ***Figure 10*** shows the results of SIM students' perception on activities in the SIM booklets.

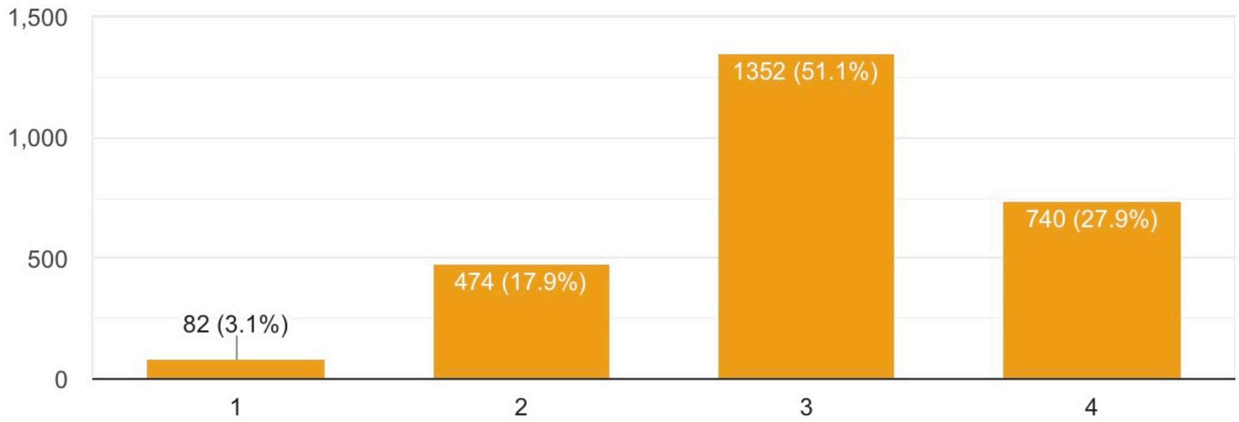


Figure 10*: Results of "Rate how did you find activities in the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 10*** the 79.0% of the SIM student respondents rated the SIM activities "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 51: Results of the SIM students' rating of SIM activities frequency distribution

```
. tabulate q25

        q25 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         82        3.10        3.10
          2 |        474       17.90       21.00
          3 |      1,352       51.06       72.05
          4 |        740       27.95      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 79.0% chose "effective" or "extremely effective."

Table 52: SIM students' rating of SIM activities frequency distribution, by age group

```
. tabulate age_group q25
```

Age_Group	q25 1	2	3	4	Total
(10-14)	30	174	630	352	1,186
(15-19)	44	200	502	237	983
(20-24)	4	21	33	12	70
(5-9)	4	79	187	139	409
Total	82	474	1,352	740	2,648

Looking at students' rating of SIM activities by age group, it shows consistently that all age groups have mode 3, which is "effective."

Table 53: SIM students' rating of SIM activities frequency distribution, by key stage

```
. tabulate key_stage q25
```

Key_Stage	q25 1	2	3	4	Total
Key Stage I	6	93	218	164	481
Key Stage II	11	100	427	245	783
Key Stage III	19	77	239	106	441
Key Stage IV	31	121	296	123	571
Key Stage V	15	83	172	102	372
Total	82	474	1,352	740	2,648

Similarly, looking at students' rating of SIM activities by key stage, it shows consistently that all key stages have mode 3, which is "effective."

Table 54: SIM students' rating of SIM activities frequency distribution, by school type

```
. tabulate school q25
```

School	q25 1	2	3	4	Total
ECR	0	5	24	39	68
HSS	52	202	453	245	952
LSS	8	32	96	40	176
MSS	14	116	377	132	639
PS	8	119	402	284	813
Total	82	474	1,352	740	2,648

Likewise, looking at students' rating of SIM activities by school type, it shows that majority of school types such as HSS, MSS, LSS, and PS rated SIM activities as "effective" with mode of 3 while interestingly ECR rated SIM activities as "extremely effective" with mode of 4. It seems lower classes appreciated activities more.

Table 55: Result of the SIM students' rating of SIM activities median calculation

```
. tabstat q25, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q25 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM activities "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 56: Result of the SIM students' measure of consensus on SIM activities rating

```
. cns q25 , min(1) max(4)

Consensus Measure for q25
Cns(X) = .6741322
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM activities rating of SIM students, it is 0.6741.

Inferential Analysis – Statistical Significance Testing

Table 57: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q25 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     2092     2890802     1753638
    Negative |      556      616474     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.384e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.410e+09

H0: q25 = 2.5
         z = 30.287
Prob > |z| = 0.0000
```

We have seen that the 79.0% of SIM students surveyed think that SIM activities was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 30.287, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 30.287 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 30.287/SQRT(2648) = 0.59. This, according to Bartz (1999) is moderate effect size.

Gender difference in SIM students' rating of SIM activities

Table 58: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q25, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1921823     1904631
        Male |     1210     1585453     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -61720724
                     ----------
Adjusted variance     3.224e+08

H0: q25(gender==Female) = q25(gender==Male)
         z =  0.958
Prob > |z| = 0.3383
```

There is no evidence for statistically significant difference between SIM activities rating between female students and male students (p-value = 0.3383 > alpha = 0.05), which means both girls and boys found SIM activities equally effective.

Evidence on SIM Students' Perception of SIM Activities

There is statistically significant evidence (p = 0.0000) that the majority 79.0% of SIM students, both girls and boys, found activities in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 30.287, p = 0.0000, with a moderate effect size (*r* = 0.59).

Effectiveness of SIM Learning

Analyzing Students' Perception on SIM Learning in Increasing Knowledge

The Ministry of Education was interested to know how SIM students found SIM learning in increasing their knowledge. To investigate this, ***Figure 11*** shows the results of SIM students' perception on increasing their knowledge during SIM learning in comparison to classroom learning.

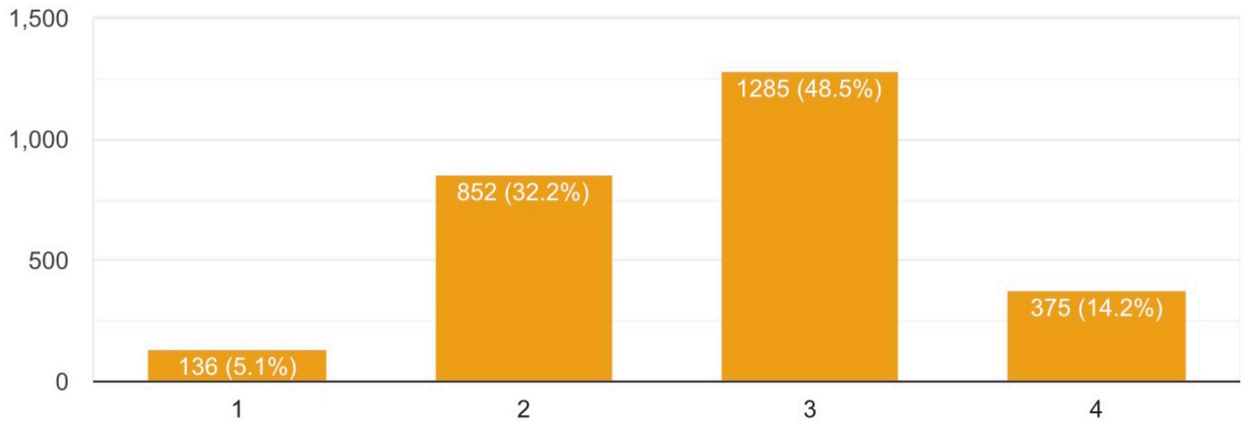


***Figure 11**: Results of "Rate the effectiveness of SIM-learning in terms of increasing knowledge" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 11*** the 62.7% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in increasing their knowledge in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 59: Results of the SIM students' rating of SIM learning in increasing knowledge

```
. tabulate q7

         q7 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        136        5.14        5.14
          2 |        852       32.18       37.31
          3 |      1,285       48.53       85.84
          4 |        375       14.16      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 62.7% chose "effective" or "extremely effective" for SIM learning in increasing their knowledge.

Table 60: SIM students' rating of SIM learning in increasing knowledge, by age group

```
. tabulate age_group q7
```

Age_Group	q7: 1	2	3	4	Total
(10-14)	36	355	623	172	1,186
(15-19)	78	323	457	125	983
(20-24)	9	21	31	9	70
(5-9)	13	153	174	69	409
Total	136	852	1,285	375	2,648

Looking at students' rating of SIM learning in increasing knowledge, by age group, it shows that in all age groups the mode is 3, which is "effective."

Table 61: SIM students' rating of SIM learning in increasing knowledge, by key stage

```
. tabulate key_stage q7
```

Key_Stage	q7: 1	2	3	4	Total
Key Stage I	14	176	213	78	481
Key Stage II	17	215	429	122	783
Key Stage III	24	158	207	52	441
Key Stage IV	54	207	249	61	571
Key Stage V	27	96	187	62	372
Total	136	852	1,285	375	2,648

Similarly, looking at students' rating of SIM learning in increasing knowledge, by key stage, it shows that consistently in all key stages, the mode is 3, which is "effective."

Table 62: SIM students' rating of SIM learning in increasing knowledge, by school type

```
. tabulate school q7
```

School	q7: 1	2	3	4	Total
ECR	3	23	28	14	68
HSS	69	312	434	137	952
LSS	4	60	87	25	176
MSS	45	238	288	68	639
PS	15	219	448	131	813
Total	136	852	1,285	375	2,648

Likewise, looking at students' rating of SIM learning in increasing knowledge, by school type, it shows that consistently in all school types, the mode is 3, which is "effective."

Table 63: Median of the SIM students' rating of SIM learning in increasing knowledge

```
. tabstat q7, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q7 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM learning "effective" or "extremely effective" in increasing their knowledge.

Descriptive Analysis - Measure of Dispersion

Table 64: SIM students' measure of consensus on SIM learning in increasing knowledge

```
. cns q7 , min(1) max(4)

Consensus Measure for q7
Cns(X) = .62691808
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in increasing knowledge, it is 0.6269.

Inferential Analysis – Statistical Significance Testing

Table 65: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q7 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1660     2271040     1753638
    Negative |      988     1236236     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -2.061e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.342e+09

H0: q7 = 2.5
         z = 14.123
Prob > |z| = 0.0000
```

We have seen that the 62.7% of SIM students surveyed think that SIM learning was effective or extremely effective in increasing their knowledge. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.123, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 14.123 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 14.123/SQRT(2648) = 0.27. This, according to Bartz (1999) is low effect size.

Gender difference in SIM students' perception of SIM learning in increasing knowledge

Table 66: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438   1920112.5     1904631
        Male |     1210   1587163.5     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -57830512
                     ----------
Adjusted variance     3.263e+08

H0: q7(gender==Female) = q7(gender==Male)
         z =  0.857
Prob > |z| = 0.3914
```

There is no evidence for statistically significant difference between female students and male students (p-value = 0.3914 > alpha = 0.05) on perception of SIM learning in increasing their knowledge, which means girls and boys rated similar on SIM learning effectiveness in increasing their knowledge.

Evidence on SIM Students' Perception of SIM Learning in Increasing Knowledge

There is statistically significant evidence (p = 0.0000) that the majority 62.7% of SIM students, both girls and boys, found SIM learning effective in increasing their knowledge. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.123, p = 0.0000, with a low effect size (*r* = 0.27).

Analyzing Students' Perception on SIM Learning in Increasing Skills

The Ministry of Education was interested to know how SIM students found SIM learning in increasing their skills. To investigate this, ***Figure 12*** shows the results of SIM students' perception on increasing their skills during SIM learning in comparison to classroom learning.

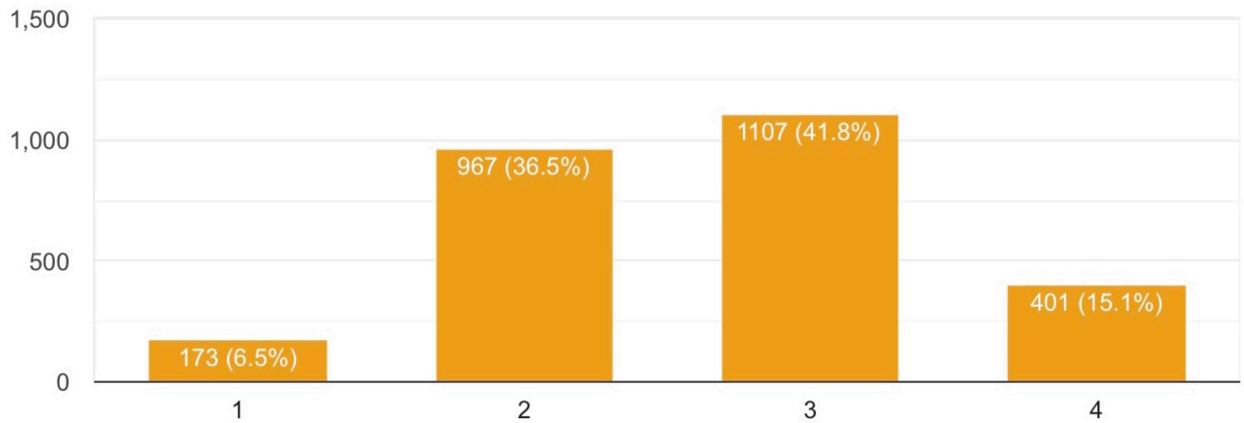


Figure 12*: Results of "Rate the effectiveness of SIM-learning in terms of increasing skills" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 12*** the 56.9% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in increasing their skills in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 67: Results of the SIM students' rating of SIM learning in increasing skills

```
. tabulate q8

         q8 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        173        6.53        6.53
          2 |        967       36.52       43.05
          3 |      1,107       41.81       84.86
          4 |        401       15.14      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 56.9% chose "effective" or "extremely effective" for SIM learning in increasing their skills.

Table 68: SIM students' rating of SIM learning in increasing skills, by age group

```
. tabulate age_group q8
```

Age_Group	q8 1	2	3	4	Total
(10-14)	50	419	553	164	1,186
(15-19)	101	334	395	153	983
(20-24)	9	23	21	17	70
(5-9)	13	191	138	67	409
Total	173	967	1,107	401	2,648

Looking at students' rating of SIM learning in increasing skills, by age group, it shows that in adolescent age groups of 10-14 and 15-19, the mode is 3, which is "effective." Interestingly, the youngest age group of 5-9 year old found SIM learning "ineffective" in increasing skills as they have mode of 2. Similarly, the oldest age group of 20-24 also rated SIM learning "ineffective" in increasing skills, but it is marginally and does not seem significant.

Table 69: SIM students' rating of SIM learning in increasing skills, by key stage

```
. tabulate key_stage q8
```

Key_Stage	q8 1	2	3	4	Total
Key Stage I	17	228	157	79	481
Key Stage II	22	271	375	115	783
Key Stage III	33	146	208	54	441
Key Stage IV	72	209	213	77	571
Key Stage V	29	113	154	76	372
Total	173	967	1,107	401	2,648

Similarly, looking at students' rating of SIM learning in increasing skills, by key stage, it shows that majority of the key stages except key stage I, have their mode as 3, which is "effective." The key stage I has mode as 2, which is "ineffective." Similar to rating of 5-9 year old, it seems the students in key stage I or class PP to III struggled with learning skills during SIM learning.

Table 70: SIM students' rating of SIM learning in increasing skills, by school type

```
. tabulate school q8

           |                        q8
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         2         31         20         15 |        68
       HSS |        91        329        379        153 |       952
       LSS |         5         67         78         26 |       176
       MSS |        56        242        272         69 |       639
        PS |        19        298        358        138 |       813
-----------+--------------------------------------------+----------
     Total |       173        967      1,107        401 |     2,648
```

Looking at students' rating of SIM learning in increasing skills, by school type, it shows that all school types except ECR have the mode as 3, which is "effective." Consistent with 5-9 year old age group and students in key stage I, ECR also rated SIM learning "ineffective" in increasing skills as it has its mode as 2. This clearly shows younger children struggled to learn skills during SIM learning.

Table 71: Median of the SIM students' rating of SIM learning in increasing skills

```
. tabstat q8, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q8 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM learning "effective" or "extremely effective" in increasing their skills.

Descriptive Analysis - Measure of Dispersion

Table 72: SIM students' measure of consensus on SIM learning in increasing skills

```
. cns q8 , min(1) max(4)

Consensus Measure for q8
Cns(X) = .59101599
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in increasing skills, it is 0.5910.

Inferential Analysis – Statistical Significance Testing

Table 73: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1508     2095474     1753638
    Negative |     1140     1411802     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.898e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.358e+09

H0: q8 = 2.5
         z =  9.275
Prob > |z| = 0.0000
```

We have seen that the 56.9% of SIM students surveyed think that SIM learning was effective or extremely effective in increasing their skills. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 9.275, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 9.275 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 9.275/SQRT(2648) = 0.18. This, according to Bartz (1999) is very low effect size.

Gender difference in SIM students' perception of SIM learning in increasing skills

Table 74: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q8, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1944360     1904631
        Male |     1210     1562916     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -48209500
                     ----------
Adjusted variance     3.359e+08

H0: q8(gender==Female) = q8(gender==Male)
         z =  2.168
Prob > |z| = 0.0302
```

There is evidence for statistically significant difference between female students and male students (p-value = 0.0302 < alpha = 0.05) on perception of SIM learning in increasing their skills, which means girls and boys rated differently on SIM learning effectiveness in increasing their skills. The positive z-score shows that in the population the female students rated skills in SIM learning higher than rating by male students. However, the difference or effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 2.168/SQRT(2648) = 0.04. This, according to Bartz (1999), is very low effect size.

Table 75: SIM students' rating of SIM learning in increasing skills, by gender

```
. tabulate gender q8

           |                          q8
    Gender |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
    Female |        94        492        627        225 |     1,438
      Male |        79        475        480        176 |     1,210
-----------+--------------------------------------------+----------
     Total |       173        967      1,107        401 |     2,648
```

Looking at students' rating of SIM learning in increasing skills, by gender, it shows that the female group has the mode clearly as 3, which is "effective" whereas the male group has mode marginally as 3 which is "effective" with frequency of 480 against next highest frequency of 475 for rating of 2, which is "ineffective."

Evidence on SIM Students' Perception of SIM Learning in Increasing Skills

There is statistically significant evidence ($p = 0.0000$) that the majority 56.9% of SIM students, with very low but significant difference between girls and boys, found SIM learning effective in increasing their skills. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 9.275$, $p = 0.0000$, with a very low effect size ($r = 0.18$).

Analyzing Students' Perception on SIM Learning in Imparting Values

The Ministry of Education was interested to know how SIM students found SIM learning in imparting values. To investigate this, ***Figure 13*** shows the results of SIM students' perception on imparting values during SIM learning in comparison to classroom learning.

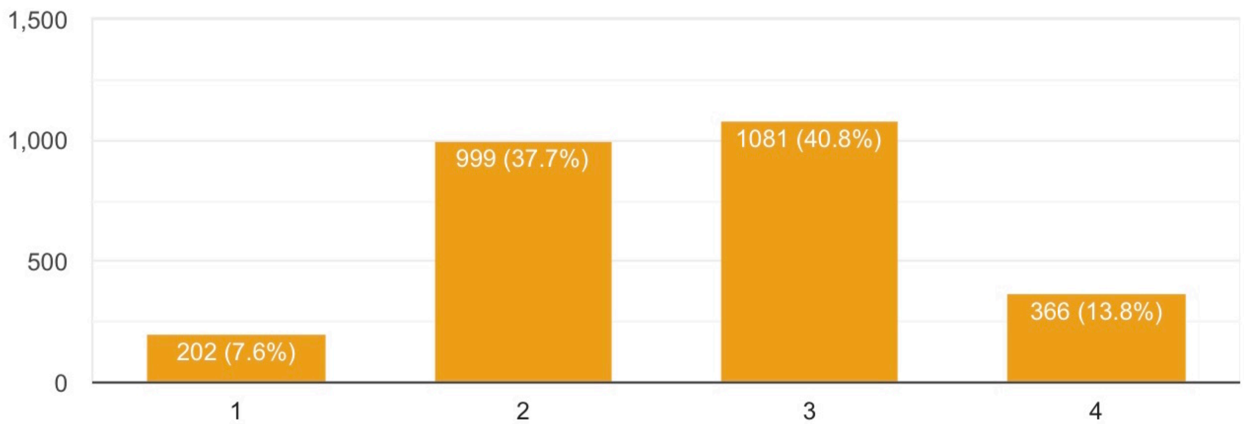


Figure 13*: Results of "Rate the effectiveness of SIM-learning in terms of imparting values" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 13*** the 54.6% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in imparting values in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 76: Results of the SIM students' rating of SIM learning in imparting values

```
. tabulate q9

         q9 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        202        7.63        7.63
          2 |        999       37.73       45.35
          3 |      1,081       40.82       86.18
          4 |        366       13.82      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 54.6% chose "effective" or "extremely effective" for SIM learning in imparting values.

Table 77: SIM students' rating of SIM learning in imparting values, by age group

```
. tabulate age_group q9

           |                    q9
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (10-14) |        73        432        504        177 |     1,186
   (15-19) |       106        378        383        116 |       983
   (20-24) |         7         25         28         10 |        70
     (5-9) |        16        164        166         63 |       409
-----------+--------------------------------------------+----------
     Total |       202        999      1,081        366 |     2,648
```

Looking at students' rating of SIM learning in imparting values, by age group, it shows that in all age groups, the mode is 3, which is "effective."

Table 78: SIM students' rating of SIM learning in imparting values, by key stage

```
. tabulate key_stage q9

              |                    q9
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |        21        199        188         73 |       481
 Key Stage II |        37        282        338        126 |       783
Key Stage III |        50        149        182         60 |       441
 Key Stage IV |        71        227        210         63 |       571
  Key Stage V |        23        142        163         44 |       372
--------------+--------------------------------------------+----------
        Total |       202        999      1,081        366 |     2,648
```

Similarly, looking at students' rating of SIM learning in imparting values, by key stage, it shows that majority of the key stages except key stages I and IV have mode as 3, which is "effective." However, the key stages I and IV have mode as 2, which is "ineffective."

Table 79: SIM students' rating of SIM learning in imparting values, by school type

```
. tabulate school q9

           |                    q9
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1         30         27         10 |        68
       HSS |        85        354        395        118 |       952
       LSS |         7         61         87         21 |       176
       MSS |        77        266        221         75 |       639
        PS |        32        288        351        142 |       813
-----------+--------------------------------------------+----------
     Total |       202        999      1,081        366 |     2,648
```

Students' rating of SIM learning in imparting values by school type is also mixed. The majority of the school types have mode as 3, which is "effective." However, ECR and MSS have mode as 2, which is "ineffective."

Table 80: Median of the SIM students' rating of SIM learning in imparting values

```
. tabstat q9, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q9 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM learning "effective" or "extremely effective" in imparting values to them.

Descriptive Analysis - Measure of Dispersion

Table 81: SIM students' measure of consensus on SIM learning in imparting values

```
. cns q9 , min(1) max(4)

Consensus Measure for q9
Cns(X) = .58541492
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in imparting values, it is 0.5854.

Inferential Analysis – Statistical Significance Testing

Table 82: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q9 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1447   1990187.5     1753638
    Negative |     1201   1517088.5     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.913e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.357e+09

H0: q9 = 2.5
         z =  6.422
Prob > |z| = 0.0000
```

We have seen that the 54.6% of SIM students surveyed think that SIM learning was effective or extremely effective in imparting values. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.422, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 6.422 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 6.422/SQRT(2648) = 0.13. This, according to Bartz (1999) is very low effect size.

Gender difference in SIM students' perception of SIM learning in imparting values

Table 83: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q9, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1960638     1904631
        Male |     1210     1546638     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -47941105
                     ----------
Adjusted variance     3.362e+08

H0: q9(gender==Female) = q9(gender==Male)
         z =  3.055
Prob > |z| = 0.0023
```

There is evidence for statistically significant difference between female students and male students (p-value = 0.0023 < alpha = 0.05) on perception of SIM learning in imparting values, which means girls and boys rated differently on SIM learning effectiveness in imparting values. The positive z-score shows that in the population the female students rated values in SIM learning higher than rating by male students. The difference or effect size (Rosenthal, 1991), r = test statistic/square root of sample size, is 3.055/SQRT(2648) = 0.06. This, according to Bartz (1999), is very low effect size.

Table 84: SIM students' rating of SIM learning in imparting values, by gender

```
. tabulate gender q9

           |                         q9
    Gender |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
    Female |        92        523        615        208 |     1,438
      Male |       110        476        466        158 |     1,210
-----------+--------------------------------------------+----------
     Total |       202        999      1,081        366 |     2,648
```

Looking at students' rating of SIM learning in imparting values, by gender, it shows that the female group has the mode as 3, which is "effective" whereas the male group has mode as 2 which is "ineffective." In other words, the female students rated SIM learning effective for imparting values but the male students rated SIM learning ineffective for imparting values.

Evidence on SIM Students' Perception of SIM Learning in Imparting Values

There is statistically significant evidence (p = 0.0000) that the majority 54.6% of SIM students found SIM learning effective in imparting values. However, there is a very low but significant difference between girls and boys where girls found SIM learning effective in imparting values but boys found it ineffective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.422, p = 0.0000, with a very low effect size (*r* = 0.13).

Analyzing Students' Perception on SIM Learning in Improving Attitudes

The Ministry of Education was interested to know how SIM students found SIM learning in improving students' attitudes. To investigate this, ***Figure 14*** shows the results of SIM students' perception on improving attitudes during SIM learning in comparison to classroom learning.

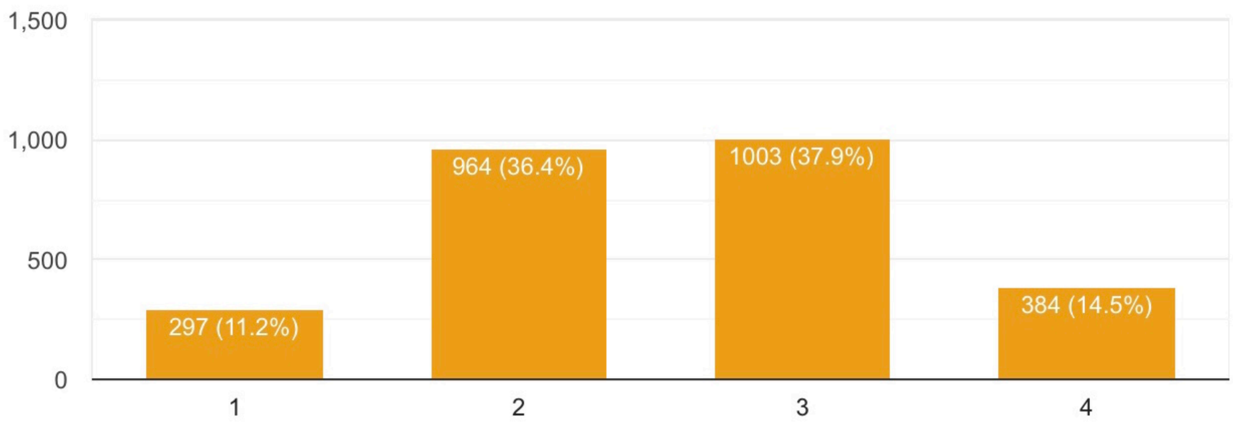


***Figure 14**: Results of "Rate the effectiveness of SIM-learning in terms of improving attitudes" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 14*** the 52.4% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in improving attitudes in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 85: Results of the SIM students' rating of SIM learning in improving attitudes

```
. tabulate q10

        q10 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        297       11.22       11.22
          2 |        964       36.40       47.62
          3 |      1,003       37.88       85.50
          4 |        384       14.50      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 52.4% chose "effective" or "extremely effective" for SIM learning in improving attitudes.

Table 86: SIM students' rating of SIM learning in improving attitudes, by age group

```
. tabulate age_group q10
```

Age_Group	q10: 1	2	3	4	Total
(10-14)	106	416	476	188	1,186
(15-19)	151	359	344	129	983
(20-24)	13	20	25	12	70
(5-9)	27	169	158	55	409
Total	297	964	1,003	384	2,648

Looking at students' rating of SIM learning in improving attitudes, by age group, it shows that the results are mixed. The age groups 10-14 and 20-24 have the mode as 3, which is "effective". But the age groups 5-9 and 15-19 have the mode as 2, which is "ineffective."

Table 87: SIM students' rating of SIM learning in improving attitudes, by key stage

```
. tabulate key_stage q10
```

Key_Stage	q10: 1	2	3	4	Total
Key Stage I	28	198	186	69	481
Key Stage II	55	270	332	126	783
Key Stage III	68	152	155	66	441
Key Stage IV	101	207	185	78	571
Key Stage V	45	137	145	45	372
Total	297	964	1,003	384	2,648

Looking at students' rating of SIM learning in improving attitudes, by key stage, it shows that the results are mixed too. The key stages II, III and V have the mode as 3, which is "effective". But the key stages I and IV have the mode as 2, which is "ineffective."

Table 88: SIM students' rating of SIM learning in improving attitudes, by school type

```
. tabulate school q10

           |                    q10
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         2         29         24         13 |        68
       HSS |       127        340        360        125 |       952
       LSS |         6         64         73         33 |       176
       MSS |       117        242        181         99 |       639
        PS |        45        289        365        114 |       813
-----------+--------------------------------------------+----------
     Total |       297        964      1,003        384 |     2,648
```

Like by key stage, students' rating of SIM learning in improving attitudes by school type is also mixed. The majority of the school types have mode as 3, which is "effective." However, ECR and MSS have mode as 2, which is "ineffective."

Table 89: Median of the SIM students' rating of SIM learning in improving attitudes

```
. tabstat q10, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q10 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM learning "effective" or "extremely effective" in improving attitudes.

Descriptive Analysis - Measure of Dispersion

Table 90: SIM students' measure of consensus on SIM learning in improving attitudes

```
. cns q10 , min(1) max(4)

Consensus Measure for q10
Cns(X) = .5491688
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in improving attitudes, it is 0.5492.

Inferential Analysis – Statistical Significance Testing

Table 91: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q10 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1387     1873224     1753638
    Negative |     1261     1634052     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.651e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.383e+09

H0: q10 = 2.5
         z =  3.216
Prob > |z| = 0.0013
```

We have seen that the 52.4% of SIM students surveyed think that SIM learning was effective or extremely effective in improving attitudes. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0013, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 3.216, p = 0.0013. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 3.216 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 3.216/SQRT(2648) = 0.06. This, according to Bartz (1999) is very low effect size.

Gender difference in SIM students' perception of SIM learning in improving attitudes

Table 92: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q10, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438   1931658.5     1904631
        Male |     1210   1575617.5     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -41118670
                     ----------
Adjusted variance     3.430e+08

H0: q10(gender==Female) = q10(gender==Male)
         z =  1.459
Prob > |z| = 0.1445
```

There is no evidence for statistically significant difference between female students and male students (p-value = 0.1445 > alpha = 0.05) on perception of SIM learning in improving attitudes, which means girls and boys rated similar on SIM learning effectiveness in improving attitudes.

Table 93: SIM students' rating of SIM learning in improving attitudes, by gender

```
. tabulate gender q10

           |                      q10
    Gender |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
    Female |       163        502        552        221 |     1,438
      Male |       134        462        451        163 |     1,210
-----------+--------------------------------------------+----------
     Total |       297        964      1,003        384 |     2,648
```

Looking at students' rating of SIM learning in improving attitudes, by gender, it shows that the female group has the mode as 3, which is "effective" whereas the male group has mode as 2 which is "ineffective." However, two-sample Wilcoxon rank-sum (Mann-Whitney) test indicated the difference is not statistically significant (p-value = 0.1445).

Evidence on SIM Students' Perception of SIM Learning in Improving Attitudes

There is statistically significant evidence ($p = 0.0013$) that the majority 52.4% of SIM students found SIM learning effective in improving attitudes. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 3.216$, $p = 0.0013$, with a very low effect size ($r = 0.06$).

Analyzing Students' Perception on SIM Learning in Understanding English

The Ministry of Education was interested to know how SIM students found SIM learning in understanding English. To investigate this, ***Figure 15*** shows the results of SIM students' perception on understanding English during SIM learning in comparison to classroom learning.

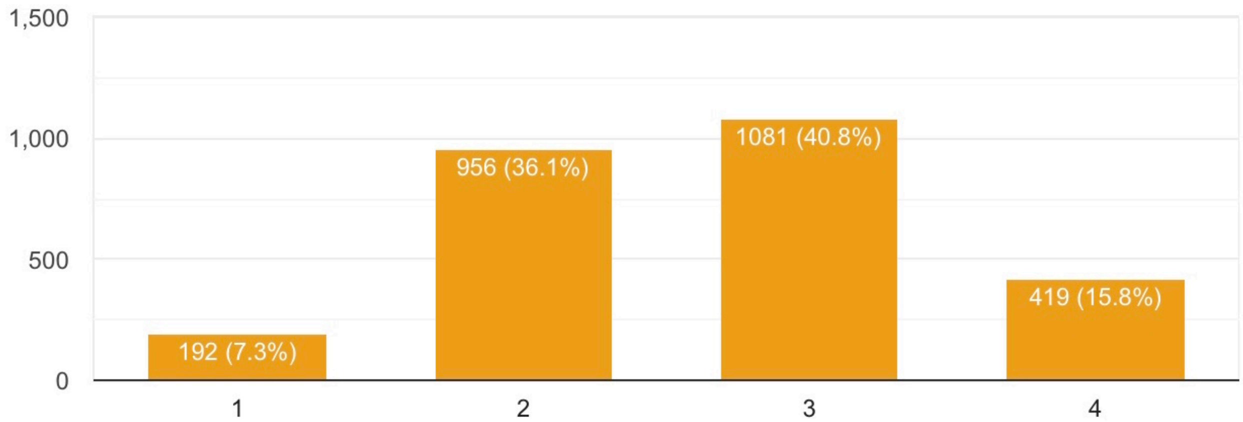


Figure 15*: Results of "Rate the effectiveness of SIM-learning in terms of understanding English subject" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 15*** the 56.6% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in understanding English subject in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 94: Results of the SIM students' rating of SIM learning in understanding English subject

```
. tabulate q11

        q11 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        192        7.25        7.25
          2 |        956       36.10       43.35
          3 |      1,081       40.82       84.18
          4 |        419       15.82      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 56.6% chose "effective" or "extremely effective" for SIM learning in understanding English subject.

Table 95: SIM students' rating of SIM learning in understanding English, by age group

```
. tabulate age_group q11
```

Age_Group	q11 1	2	3	4	Total
(10-14)	56	410	541	179	1,186
(15-19)	93	333	390	167	983
(20-24)	13	19	25	13	70
(5-9)	30	194	125	60	409
Total	192	956	1,081	419	2,648

Looking at students' rating of SIM learning in understanding English subject, by age group, it shows that the majority of the age groups except 5-9 age group have the mode as 3, which is "effective". But the 5-9 age group has the mode as 2, which is "ineffective." It seems the younger children had difficulty in understanding English during SIM learning.

Table 96: SIM students' rating of SIM learning in understanding English, by key stage

```
. tabulate key_stage q11
```

Key_Stage	q11 1	2	3	4	Total
Key Stage I	33	224	158	66	481
Key Stage II	24	277	361	121	783
Key Stage III	44	150	187	60	441
Key Stage IV	63	198	223	87	571
Key Stage V	28	107	152	85	372
Total	192	956	1,081	419	2,648

Similar to age group ratings, looking at students' rating of SIM learning in understanding English subject by key stage, it shows that the majority of the key stages except key stage I have the mode as 3, which is "effective". But the key stage I has the mode as 2, which is "ineffective." It seems the children in lower classes had difficulty in understanding English during SIM learning.

Table 97: SIM students' rating of SIM learning in understanding English, by school type

```
. tabulate school q11

           |                        q11
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         3         35         23          7 |        68
       HSS |        91        327        363        171 |       952
       LSS |        11         64         77         24 |       176
       MSS |        64        217        281         77 |       639
        PS |        23        313        337        140 |       813
-----------+--------------------------------------------+----------
     Total |       192        956      1,081        419 |     2,648
```

Looking at students' rating of SIM learning in understanding English by school type, it shows the majority of the school types except ECR have mode as 3, which is "effective." However, ECR has mode as 2, which is "ineffective."

Table 98: Median of the SIM students' rating of SIM learning in understanding English

```
. tabstat q11, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q11 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM learning "effective" or "extremely effective" in understanding English.

Descriptive Analysis - Measure of Dispersion

Table 99: SIM students' measure of consensus on SIM learning in understanding English

```
. cns q11 , min(1) max(4)

Consensus Measure for q11
Cns(X) = .57987788
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in understanding English, it is 0.5799.

Inferential Analysis – Statistical Significance Testing

Table 100: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q11 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1500     2083256     1753638
    Negative |     1148     1424020     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.808e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.367e+09

H0: q11 = 2.5
         z =  8.914
Prob > |z| = 0.0000
```

We have seen that the 56.6% of SIM students surveyed think that SIM learning was effective or extremely effective in understanding English. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.914, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 8.914 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 8.914/SQRT(2648) = 0.17. This, according to Bartz (1999) is very low effect size.

Gender difference in SIM students' perception of SIM learning in understanding English

Table 101: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q11, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1956133     1904631
        Male |     1210     1551143     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -45874245
                     ----------
Adjusted variance     3.382e+08

H0: q11(gender==Female) = q11(gender==Male)
         z =  2.800
Prob > |z| = 0.0051
```

There is evidence for statistically significant difference between female students and male students (p-value = 0.0051 < alpha = 0.05) on perception of SIM learning in understanding English, which means girls and boys rated differently on SIM learning effectiveness in understanding English subject. The difference or effect size (Rosenthal, 1991), r = test statistic/square root of sample size, is 2.800/SQRT(2648) = 0.05. This, according to Bartz (1999), is very low effect size.

Table 102: SIM students' rating of SIM learning in understanding English, by gender

```
. tabulate gender q11

           |                         q11
    Gender |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
    Female |        89        509        590        250 |     1,438
      Male |       103        447        491        169 |     1,210
-----------+--------------------------------------------+----------
     Total |       192        956      1,081        419 |     2,648
```

Looking at students' rating of SIM learning in understanding English, by gender, it shows that the female group has the mode as 3, which is "effective" and the male group also has mode as 3 which is “effective.” However, two-sample Wilcoxon rank-sum (Mann-Whitney) test indicated that there is a statistically significant difference between ratings of female students and male students (p-value = 0.0051). The girls have rated understanding English during SIM learning marginally higher than ratings by boys.

Evidence on SIM Students’ Perception of SIM Learning in Understanding English

There is statistically significant evidence ($p = 0.0000$) that the majority 56.6% of SIM students found SIM learning effective in understanding English subject. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 8.914$, $p = 0.0000$, with a very low effect size ($r = 0.17$).

Analyzing Students' Perception on SIM Learning in Understanding Mathematics

The Ministry of Education was interested to know how SIM students found SIM learning in understanding Mathematics. To investigate this, ***Figure 16*** shows the results of SIM students' perception on understanding Mathematics during SIM learning in comparison to classroom learning.

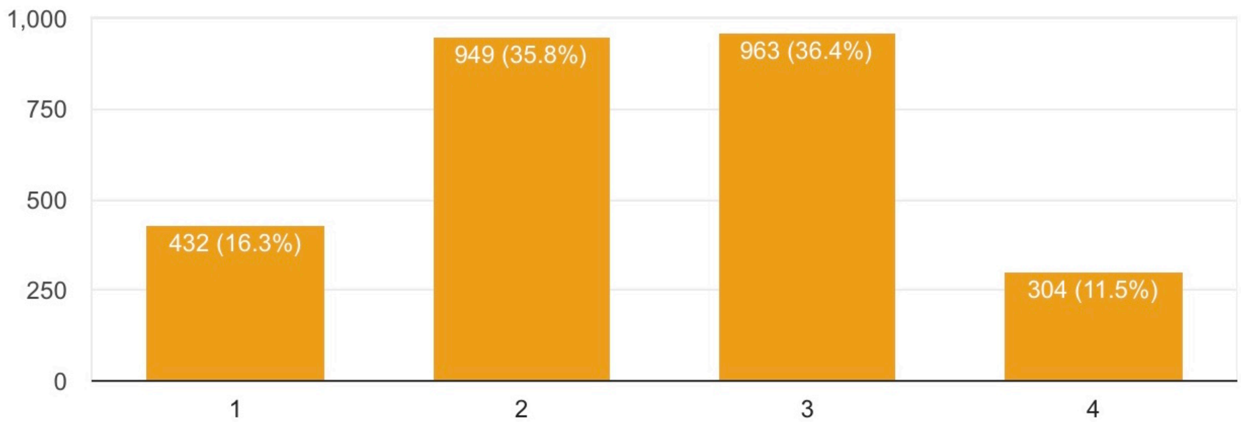


Figure 16*: Results of "Rate the effectiveness of SIM-learning in terms of understanding Mathematics subject" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 16*** only 47.9% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in understanding Mathematics subject in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 103: Results of the SIM students' rating of SIM learning in understanding Mathematics subject

```
. tabulate q12

        q12 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        432       16.31       16.31
          2 |        949       35.84       52.15
          3 |        963       36.37       88.52
          4 |        304       11.48      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 47.9% chose "effective" or "extremely effective" for SIM learning in understanding Mathematics.

Table 104: SIM students' rating of SIM learning in understanding Mathematics, by age group

```
. tabulate age_group q12

           |                         q12
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (10-14) |       106        404        519        157 |     1,186
   (15-19) |       278        343        280         82 |       983
   (20-24) |        23         23         15          9 |        70
     (5-9) |        25        179        149         56 |       409
-----------+--------------------------------------------+----------
     Total |       432        949        963        304 |     2,648
```

Looking at students' rating of SIM learning in understanding Mathematics subject, by age group, it shows that the majority of the age groups except 10-14 have the mode as 2, which is "ineffective". But the 10-14 age group has the mode as 3, which is "effective." It seems the majority of the students had difficulty in understanding Mathematics during SIM learning.

Table 105: SIM students' rating of SIM learning in understanding Mathematics, by key stage

```
. tabulate key_stage q12

              |                         q12
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |        29        210        173         69 |       481
 Key Stage II |        46        254        378        105 |       783
Key Stage III |        78        152        163         48 |       441
 Key Stage IV |       146        215        158         52 |       571
  Key Stage V |       133        118         91         30 |       372
--------------+--------------------------------------------+----------
        Total |       432        949        963        304 |     2,648
```

Looking at students' rating of SIM learning in understanding Mathematics by key stage, it shows that the majority of the key stages except key stages II and III have the mode as 2, which is "ineffective". But the key stages II and III have the mode as 3, which is "effective."

Table 106: SIM students' rating of SIM learning in understanding Mathematics, by school type

```
. tabulate school q12

           |                       q12
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         2         35         24          7 |        68
       HSS |       258        321        274         99 |       952
       LSS |        12         72         72         20 |       176
       MSS |       116        235        231         57 |       639
        PS |        44        286        362        121 |       813
-----------+--------------------------------------------+----------
     Total |       432        949        963        304 |     2,648
```

Looking at students' rating of SIM learning in understanding Mathematics by school type, it shows the majority of the school types except PS have mode as 2, which is "ineffective." However, PS has mode as 3, which is "effective." In the case of LSS, it has grey area of bimodal, both 2 and 3.

Table 107: Median of the SIM students' rating of SIM learning in understanding Mathematics

```
. tabstat q12, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q12 |      2648         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM student respondents found SIM learning "ineffective" or "extremely ineffective" in understanding Mathematics.

Descriptive Analysis - Measure of Dispersion

Table 108: SIM students' measure of consensus on SIM learning in understanding Mathematics

```
. cns q12 , min(1) max(4)

Consensus Measure for q12
Cns(X) = .53432411
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in understanding Mathematics, it is 0.5343.

Inferential Analysis – Statistical Significance Testing

Table 109: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q12 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1267   1614381.5     1753638
    Negative |     1381   1892894.5     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.539e+08
Adjustment for zeros          0
                     ----------
Adjusted variance     1.394e+09

H0: q12 = 2.5
         z = -3.729
Prob > |z| = 0.0002
```

We have seen that only 47.9% of SIM students surveyed think that SIM learning was effective or extremely effective in understanding Mathematics. In other words, 52.1% majority of the students surveyed think SIM learning was ineffective or extremely ineffective in understanding Mathematics. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = “ineffective” and 3 = “effective.”

Ho: Our null hypothesis is that the SIM student population’s true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population’s true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0002, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = - 3.729, p = 0.0002. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -3.729 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -3.729/SQRT(2648) = -0.07 or 0.07 (we can ignore the negative sign). This, according to Bartz (1999) is very low effect size.

Gender difference in SIM students' perception of SIM learning in understanding Mathematics

Table 110: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q12, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438     1869098     1904631
        Male |     1210     1638178     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -38403608
                     ----------
Adjusted variance     3.457e+08

H0: q12(gender==Female) = q12(gender==Male)
         z = -1.911
Prob > |z| = 0.0560
```

There is no evidence for statistically significant difference between female students and male students (p-value = 0.0560 > alpha = 0.05) on perception of SIM learning in understanding Mathematics, which means girls and boys rated similar on SIM learning effectiveness in understanding Mathematics.

Evidence on SIM Students' Perception of SIM Learning in Understanding Mathematics

There is statistically significant evidence (p = 0.0002) that only minority 47.9% of SIM students found SIM learning effective in understanding Mathematics. In other words, the majority 52.1% of SIM students found SIM learning ineffective in understanding Mathematics. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -3.729, p = 0.0002, with a very low effect size (*r* = 0.07).

Analyzing Students' Perception on SIM Learning in Understanding Dzongkha

The Ministry of Education was interested to know how SIM students found SIM learning in understanding Dzongkha. To investigate this, ***Figure 17*** shows the results of SIM students' perception on understanding Dzongkha during SIM learning in comparison to classroom learning.

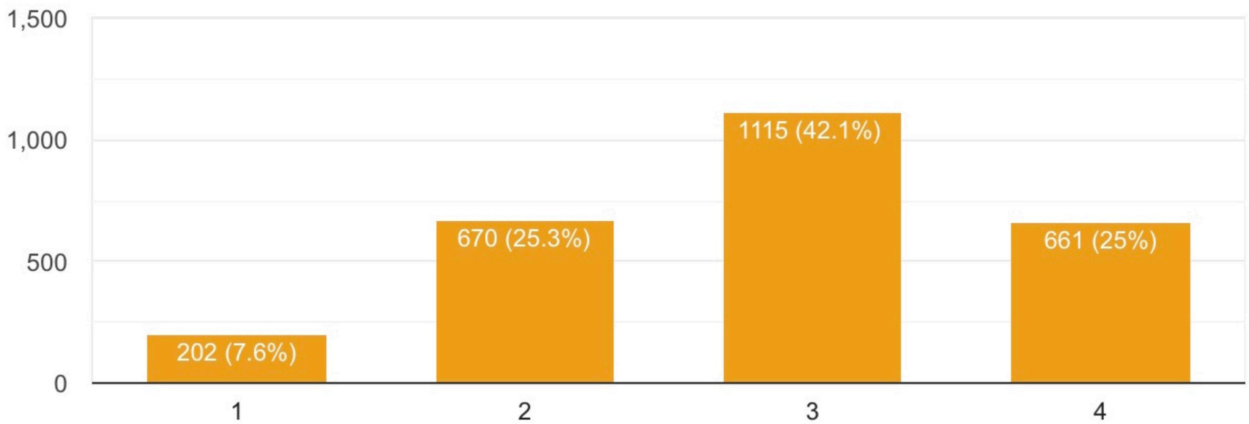


***Figure 17**: Results of "Rate the effectiveness of SIM-learning in terms of understanding Dzongkha subject" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 17*** the 67.1% of the SIM student respondents rated the SIM learning "effective" or "extremely effective" in understanding Dzongkha subject in comparison to classroom learning.

Descriptive Analysis – Measure of Central Tendency

Table 111: Results of the SIM students' rating of SIM learning in understanding Dzongkha subject

```
. tabulate q13

        q13 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        202        7.63        7.63
          2 |        670       25.30       32.93
          3 |      1,115       42.11       75.04
          4 |        661       24.96      100.00
------------+-----------------------------------
      Total |      2,648      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM student respondents of 67.1% chose "effective" or "extremely effective" for SIM learning in understanding Dzongkha subject.

Table 112: SIM students' rating of SIM learning in understanding Dzongkha, by age group

```
. tabulate age_group q13

           |                        q13
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (10-14) |        85        285        520        296 |     1,186
   (15-19) |        88        218        404        273 |       983
   (20-24) |         8         12         30         20 |        70
     (5-9) |        21        155        161         72 |       409
-----------+--------------------------------------------+----------
     Total |       202        670      1,115        661 |     2,648
```

Looking at students' rating of SIM learning in understanding Dzongkha subject, by age group, it shows that all age groups have the mode as 3, which is "effective".

Table 113: SIM students' rating of SIM learning in understanding Dzongkha, by key stage

```
. tabulate key_stage q13

              |                        q13
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |        26        178        188         89 |       481
 Key Stage II |        39        189        354        201 |       783
Key Stage III |        51         94        180        116 |       441
 Key Stage IV |        59        139        222        151 |       571
  Key Stage V |        27         70        171        104 |       372
--------------+--------------------------------------------+----------
        Total |       202        670      1,115        661 |     2,648
```

Similar to age group ratings, looking at students' rating of SIM learning in understanding Dzongkha subject by key stage, it shows that all key stages have the mode as 3, which is "effective".

Table 114: SIM students' rating of SIM learning in understanding Dzongkha, by school type

```
. tabulate school q13

           |                        q13
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         6         26         23         13 |        68
       HSS |        95        211        375        271 |       952
       LSS |        12         61         65         38 |       176
       MSS |        57        167        284        131 |       639
        PS |        32        205        368        208 |       813
-----------+--------------------------------------------+----------
     Total |       202        670      1,115        661 |     2,648
```

Looking at students' rating of SIM learning in understanding Dzongkha by school type, it shows the majority of the school types except ECR have mode as 3, which is "effective." However, ECR has mode as 2, which is "ineffective."

Table 115: Median of the SIM students' rating of SIM learning in understanding Dzongkha

```
. tabstat q13, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q13 |      2648         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM student respondents found SIM learning "effective" or "extremely effective" in understanding Dzongkha.

Descriptive Analysis - Measure of Dispersion

Table 116: SIM students' measure of consensus on SIM learning in understanding Dzongkha

```
. cns q13 , min(1) max(4)

Consensus Measure for q13
Cns(X) = .56675479
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in understanding Dzongkha, it is 0.5668.

Inferential Analysis – Statistical Significance Testing

Table 117: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q13 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |     1776     2461132     1753638
    Negative |      872     1046144     1753638
        Zero |        0           0           0
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties  -1.319e+08
Adjustment for zeros         0
                     ----------
Adjusted variance     1.416e+09

H0: q13 = 2.5
         z = 18.800
Prob > |z| = 0.0000
```

We have seen that the 67.1% of SIM students surveyed think that SIM learning was effective or extremely effective in understanding Dzongkha. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM student population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM student population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM student population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 18.800, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 18.800 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 18.800/SQRT(2648) = 0.37. This, according to Bartz (1999) is low effect size.

Gender difference in SIM students' perception of SIM learning in understanding Dzongkha

Table 118: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q13, by(gender)

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |     1438   1982740.5     1904631
        Male |     1210   1524535.5     1602645
-------------+---------------------------------
    Combined |     2648     3507276     3507276

Unadjusted variance   3.841e+08
Adjustment for ties   -41042471
                     ----------
Adjusted variance     3.431e+08

H0: q13(gender==Female) = q13(gender==Male)
         z =  4.217
Prob > |z| = 0.0000
```

There is evidence for statistically significant difference between female students and male students (p-value = 0.0000 < alpha = 0.05) on perception of SIM learning in understanding Dzongkha, which means girls and boys rated differently on SIM learning effectiveness in understanding Dzongkha subject. The difference or effect size (Rosenthal, 1991), r = test statistic/square root of sample size, is 4.217/SQRT(2648) = 0.08. This, according to Bartz (1999), is very low effect size.

Table 119: SIM students' rating of SIM learning in understanding Dzongkha, by gender

```
. tabulate gender q13

           |                         q13
    Gender |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
    Female |        89        342        613        394 |     1,438
      Male |       113        328        502        267 |     1,210
-----------+--------------------------------------------+----------
     Total |       202        670      1,115        661 |     2,648
```

Looking at students' rating of SIM learning in understanding Dzongkha, by gender, it shows that both girls and boys have the mode as 3, which is "effective." However, two-sample Wilcoxon rank-sum (Mann-Whitney) test indicated that there is a statistically significant difference between ratings of female students and male students (p-value = 0.0000). The girls have rated understanding Dzongkha during SIM learning marginally higher than ratings by boys.

Evidence on SIM Students' Perception of SIM Learning in Understanding Dzongkha

There is statistically significant evidence ($p = 0.0000$) that the majority 67.1% of SIM students found SIM learning effective in understanding Dzongkha subject. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 18.800$, $p = 0.0000$, with a low effect size ($r = 0.37$).

Advantages and Disadvantages of SIM Learning

Analyzing SIM Students' Perception of Advantages of SIM Learning

The Ministry of Education was interested to know what SIM students found as advantages of SIM learning. To investigate this, ***Figure 18*** shows the results of SIM students' perception of advantages of SIM learning.

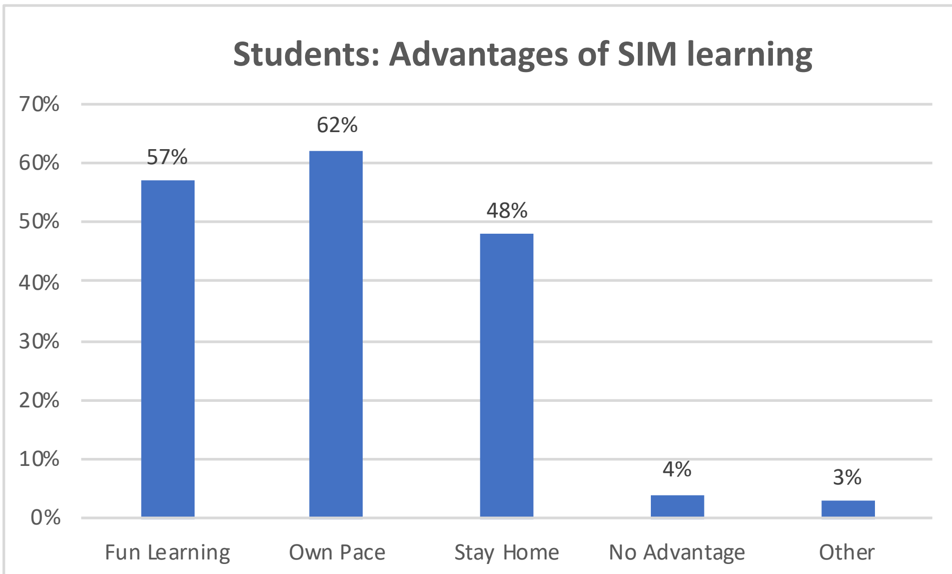


Figure 18: *Results of "What are the advantages of SIM-learning?"*

As shown in ***Figure 18,*** the SIM students found "Learning on your own pace" (62%) as the main advantage of SIM learning, followed by "Self-learning is fun" (57%) and "Ability to stay at home" (48%).

Inferential Analysis – Statistical Significance Testing through Cochran's Q Test

To test if the differences between advantages of SIM learning are significantly different we can use a Cochran's Q test.

Table 120: Results of Cochran's Q Test on Advantages of SIM Learning

```
. cochran q5_1 q5_2 q5_3 q5_4 q5_5, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
        q5_1 |   .5679758       1504
        q5_2 |   .6242447       1653
        q5_3 |   .4784743       1267
        q5_4 |   .0411631        109
        q5_5 |   .0271903         72
------------------------------------

Number of obs        =      2648
Cochran's chi2(4)    =  3604.269
Prob > chi2          =    0.0000
```

We have seen that the 62% of SIM students surveyed think that the main advantage of SIM learning was "Learning on your own pace," followed by "Self-learning is fun" (57%) and "Ability to stay at home" (48%). However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether there are differences between the proportions among the five options of advantages of SIM learning.

Ho: Our null hypothesis is that there are no differences between the proportions among the five options of advantages of SIM learning.

Ha: Our alternative hypothesis is that there are statistically significant differences between the proportions among the five options of advantages of SIM learning.

Cochran's Q test would show us how likely to have result as in our survey sample or even higher, if in the population there would be no differences. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that most likely in the population each option is not chosen equally often. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of advantages of SIM learning, χ^2 $(4, N = 2648) = 3604.269$, $p = 0.0000$.

Post-hoc test

Since there are statistically significant differences in proportions of advantages of SIM learning, we would like to know whether there is statistically significant difference between "Learning on your own pace" (62%) and "Self-learning is fun" (57%) through pairwise comparisons as these two options were majority of the SIM students' choices on advantages of SIM learning. We will use Cochran's test for pairs.

Table 121: Results of Cochran's Q post-hoc test

```
. cochran q5_1 q5_2, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+---------------------
        q5_1 |   .5679758       1504
        q5_2 |   .6242447       1653
------------------------------------

Number of obs       =      2648
Cochran's chi2(1)   =  14.83033
Prob > chi2         =    0.0001
Exact p             =    0.0001
```

A pairwise post-hoc Cochran's Q test was statistically significant for "Learning at your own pace" vs. "Self-learning is fun", χ^2 (1, N = 2648) = 14.83033, p = 0.0001. Therefore, the number one advantage of SIM learning for SIM students was "Learning at your own pace." The effect size between them η^{2} = 14.83033/2648 = 0.0056 = 0.01.

Effect Size

The test statistic is Q = 3604.269, our sample size for SIM students is 2648 and we have five options (variables) for advantages of SIM learning. Therefore, the effect size for this can be calculated by eta-squared (η^2) (Serlin, Carr, & Marascuilo, 1982).

η^2 = 3604.269/((5-1)x2648) = 0.34, which is a large effect size.

Evidence on SIM Students' Perception on Advantages of SIM Learning

There is statistically significant evidence (p = 0.0000) that the majority of SIM students found "Learning at your own pace" as the main advantage of SIM learning, followed by "Self-learning is fun". In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of advantages of SIM learning, χ^2(4, N = 2648) = 3604.269, p =0.0000, with a large effect size (η^2 = 0.34). A pairwise post-hoc Cochran test was also significant for "Learning at your own pace" vs. "Self-learning is fun" (p = .0001) but the difference (effect size) between them is very small (η^2 = 0.01).

Analyzing SIM Students' Perception on Disadvantages of SIM Learning

The Ministry of Education was interested to know what SIM students found as disadvantages of SIM learning. To investigate this, ***Figure 19*** shows the results of SIM students' perception of disadvantages of SIM learning.

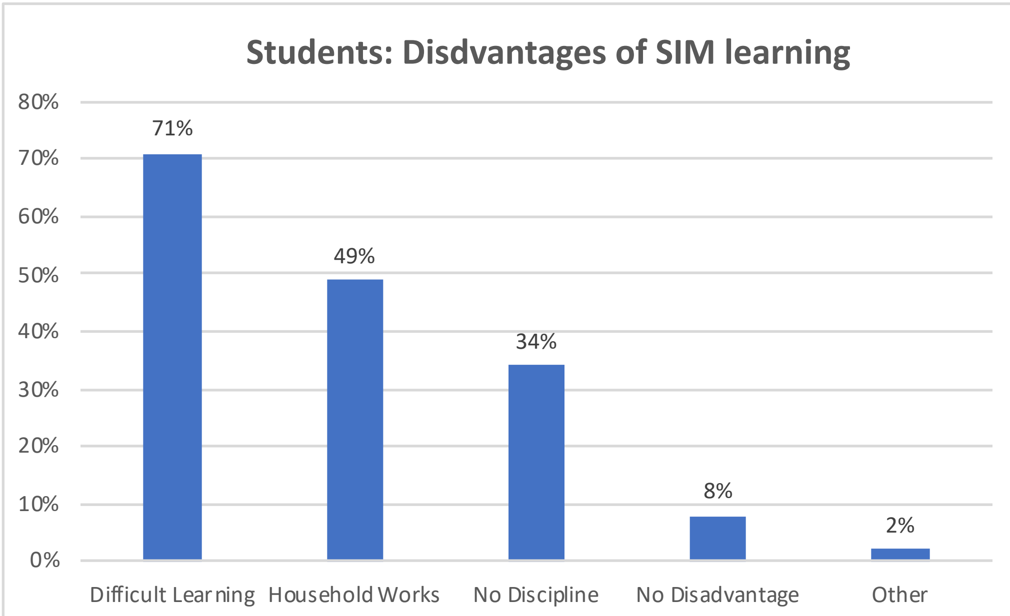


Figure 19: *Results of "What are the disadvantages of SIM-learning?"*

As shown in ***Figure 19,*** the SIM students found "Self-learning is difficult" (71%) as the main disadvantage of SIM learning, followed by "Household works at home" (49%) and "No self-discipline" (34%).

Inferential Analysis – Statistical Significance Testing through Cochran's Q Test

To test if the differences between disadvantages of SIM learning are significantly different we can use a Cochran's Q test.

Table 122: Results of Cochran's Q Test on Disadvantages of SIM Learning

```
. cochran q6_1 q6_2 q6_3 q6_4 q6_5, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
        q6_1 |   .7054381       1868
        q6_2 |   .4882931       1293
        q6_3 |   .3387462        897
        q6_4 |   .0785498        208
        q6_5 |   .0181269         48
------------------------------------

Number of obs        =       2648
Cochran's chi2(4)    =   3558.177
Prob > chi2          =     0.0000
```

We have seen that the 71% of SIM students surveyed think that the main disadvantage of SIM learning was "Self-learning is difficult," followed by "Household works at home" (49%) and "No self-discipline" (34%). However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether there are differences between the proportions among the five options of disadvantages of SIM learning.

Ho: Our null hypothesis is that there are no differences between the proportions among the five options of disadvantages of SIM learning.

Ha: Our alternative hypothesis is that there are statistically significant differences between the proportions among the five options of disadvantages of SIM learning.

Cochran's Q test would show us how likely to have result as in our survey sample or even higher, if in the population there would be no differences. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that most likely in the population each option is not chosen equally often. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of disadvantages of SIM learning, χ^2 $(4, N = 2648) = 3558.177$, $p = 0.0000$.

Post-hoc test

Since there are statistically significant differences in proportions of disadvantages of SIM learning, we would like to know whether there is statistically significant difference between

“Self-learning is difficult” (71%) and “Household works at home” (49%) through pairwise comparisons as these two options are most selected of the SIM students’ choices on disadvantages of SIM learning. We will use Cochran’s test for pairs.

Table 123: Results of Cochran’s Q post-hoc test

```
. cochran q6_1 q6_2, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
        q6_1 |   .7054381       1868
        q6_2 |   .4882931       1293
------------------------------------

Number of obs       =       2648
Cochran's chi2(1)   =   227.2337
Prob > chi2         =     0.0000
Exact p             =     0.0000
```

A pairwise post-hoc Cochran’s Q test was statistically significant for “Self-learning is difficult” vs. “Household works at home”, χ^2 (1, N = 2648) = 227.2337, p = 0.0000. Therefore, the number one disadvantage of SIM learning for SIM students was “Self-learning is difficult.” The effect size between them $\eta^2 = 227.2337/2648 = 0.0858 = 0.09$, which is a medium effect size.

Effect Size

The test statistic is Q = 3558.177, our sample size for SIM students is 2648 and we have five options (variables) for disadvantages of SIM learning. Therefore, the effect size for this can be calculated by eta-squared (η^2) (Serlin, Carr, & Marascuilo, 1982).

η^2 = 3558.177/((5-1)x2648) = 0.34, which is a large effect size.

Evidence on SIM Students’ Perception of Disadvantages of SIM Learning

There is statistically significant evidence (p = 0.0000) that the majority of SIM students found “Self-learning is difficult” as the main and only disadvantage of SIM learning. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of disadvantages of SIM learning, χ^2(4, N = 2648) = 3558.177, p =0.0000, with a large effect size (η^2 = 0.34). A pairwise post-hoc Cochran test was also significant for “Self-learning is difficult” vs. “Household works at home” (p = .0000) with a moderate effect size (η^2 = 0.09).

Effect of Household Chores on SIM Learning

Significance of Household Chores on SIM Learning: Is "Household works at home" a statistically significant disadvantage for the majority of the SIM students?

As a social norm perception, usually people think having to do household works or chores at home is a disadvantage for studying at home, especially for adolescent girls during the COVID-19 pandemic. In this SIM program assessment study, we surveyed and tested this perception too. We found only 49% of the SIM students surveyed selected "Household works at home" as a disadvantage for SIM learning. We need to test whether the majority of the SIM students in the population would select "Household works at home" as a disadvantage or not.

Ho: Our null hypothesis is that the percentage of the SIM students who selected "Household works at home" as a disadvantage is 50%.

Ha: Our alternative hypothesis is the percentage of the SIM students who selected "Household works at home" as a disadvantage is greater than 50%.

Table 124: Results of One Sample Binomial Test on Household Works

```
. bitest q6_2 = 0.50

Binomial probability test

    Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+-------------------------------------------------------------
        q6_2 |      2,648        1,293        1,324     0.50000      0.48829

  Pr(k >= 1,293)            = 0.889581  (one-sided test)
  Pr(k <= 1,293)            = 0.117923  (one-sided test)
  Pr(k <= 1,293 or k >= 1,355) = 0.235847  (two-sided test)
```

One-sided binomial test indicated that the percentage of SIM students who selected "Household works at home" as a disadvantage (N_{hw} = 1293, 49%), was not statistically significantly different from the population hypothesized value of 50%, p = 0.889581 (which is much greater than alpha = 0.05). Therefore, there is no sufficient evidence that "Household works at home" affected the majority of SIM students during SIM learning.

Gender Difference in Effect of Household Chores in SIM Learning: Is there gender difference in "Household works at home" for the SIM students?

Table 125: Results of Two-Sample Test of Proportions on Household Works, by Gender

```
. prtest q6_2, by(gender)

Two-sample test of proportions                Female: Number of obs =     1438
                                                Male: Number of obs =     1210
------------------------------------------------------------------------------
       Group |       Mean   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
      Female |   .4888734   .0131821                       .4630371    .5147098
        Male |   .4876033   .0143696                       .4594395    .5157671
-------------+----------------------------------------------------------------
        diff |   .0012701      .0195                      -.0369492    .0394895
             |  under H0:   .0195001     0.07   0.948
------------------------------------------------------------------------------
        diff = prop(Female) - prop(Male)                          z =   0.0651
    H0: diff = 0

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.5260         Pr(|Z| > |z|) = 0.9481          Pr(Z > z) = 0.4740
```

Since our SIM survey sample is large enough (N=2648) to assume normal distribution, we applied two-sample test of proportions to test whether "Household works at home" affected girls more than boys during SIM learning in times of COVID-19 pandemic. We found that there is no statistically significant evidence that girls were affected more than boys by "Household works at home" during the SIM learning, $z = 0.0651$, $p = 0.4740$ (which is greater than alpha = 0.05). Therefore, "Household works at home" was not statistically significant disadvantage for the majority of students, both boys and girls, during SIM learning.

Help Sought for SIM Learning

Analyzing SIM Students' Perception of Help Sought for SIM Learning

The Ministry of Education was interested to know if SIM students sought help during SIM learning and if so, from whom. To investigate this, ***Figure 20*** shows the results of SIM students' perception on help sought during SIM learning.

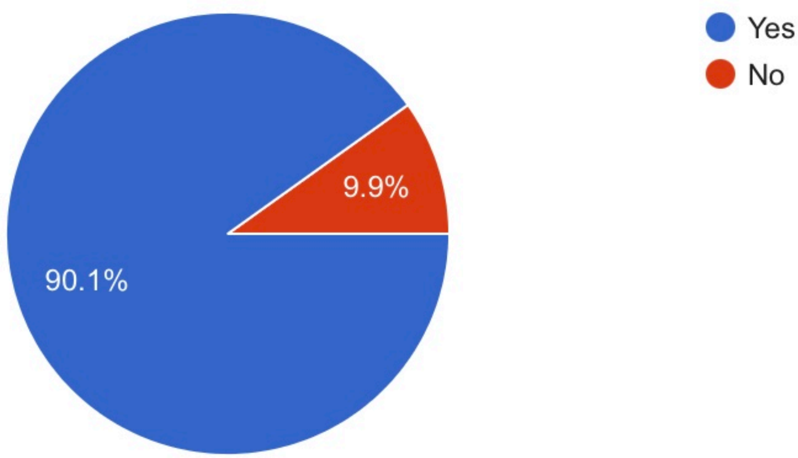


Figure 20: *Results of "Did you seek help from anyone to understand SIM lessons?"*

As shown in ***Figure 20,*** the vast 90.1% majority of SIM students said they sought help from someone to understand SIM lessons.

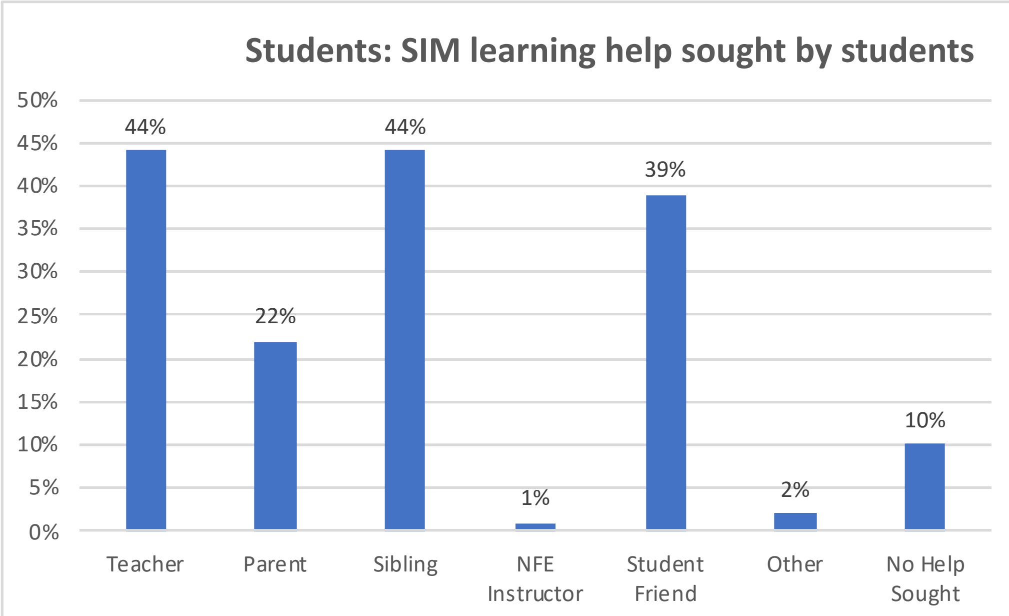


Figure 21: *Results of "From whom did you seek help to understand SIM lessons?"*

As shown in ***Figure 21,*** the SIM students mainly sought help from teachers (44%) and siblings (44%), followed by student friends (39%) and parents (22%). Against a popular belief that SIM students would seek help from NFE instructors in the rural areas, only about 1% of the SIM students actually sought help from NFE instructors. About 10% of SIM students did not seek help from anyone.

Inferential Analysis – Statistical Significance Testing

Table 126: Results of Binomial Test on Help Sought for SIM lessons

```
. bitest q28a = 0.89

Binomial probability test

    Variable |       N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
        q28a |   2,648      2,386      2,356.72     0.89000      0.90106

  Pr(k >= 2,386)            = 0.035444  (one-sided test)
  Pr(k <= 2,386)            = 0.969275  (one-sided test)
  Pr(k <= 2,327 or k >= 2,386) = 0.071594  (two-sided test)
```

A binomial test indicated that the percentage of SIM students who sought help for SIM lessons (N_{help} = 2386, 90.1%) was statistically significantly greater than the population hypothesized value of 89%, p = 0.035444 (which is less than significance level alpha = 0.05). It means at least 89% of the SIM students sought help during SIM learning.

Similarly, to test if the differences between SIM helpers are significantly different we can use a Cochran's Q test.

Table 127: Results of Cochran's Q Test on helpers of SIM lessons

```
. cochran q29_1 q29_2 q29_3 q29_4 q29_5, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
       q29_1 |    .435423       1153
       q29_2 |   .2228097        590
       q29_3 |   .4373112       1158
       q29_4 |   .0098187         26
       q29_5 |   .3882175       1028
------------------------------------

Number of obs        =       2648
Cochran's chi2(4)    =   1670.831
Prob > chi2          =     0.0000
```

We have seen that the 43.5% of SIM students surveyed said that they took help from teachers and 43.7% of SIM students said they took help from siblings, followed by 38.8% for student friends and 22.3% for parents. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether there are differences between the proportions among the five options of help for SIM lessons.

Ho: Our null hypothesis is that there are no differences between the proportions among the five options of help for SIM lessons.

Ha: Our alternative hypothesis is that there are statistically significant differences between the proportions among the five options of help for SIM lessons.

Cochran's Q test would show us how likely to have result as in our survey sample or even higher, if in the population there would be no differences. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that most likely in the population each option is not chosen equally often. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of help for SIM lessons, χ^2 (4, N = 2648) = 1670.831, p = 0.0000.

Post-hoc test

Since there are statistically significant differences in proportions of helpers for SIM lessons, we would like to know whether there is statistically significant difference between "Teacher" (43.5%) and "Sibling" (43.7%) through pairwise comparisons as these two options are most selected help options by SIM students. We will use Cochran's test for pairs.

Table 128: Results of Cochran's Q post-hoc test

```
. cochran q29_1 q29_3, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
       q29_1 |    .435423       1153
       q29_3 |   .4373112       1158
------------------------------------

Number of obs       =       2648
Cochran's chi2(1)   =   .0191571
Prob > chi2         =     0.8899
Exact p             =     0.9118
```

An exact pairwise post-hoc Cochran's Q test was not statistically significant for "Teacher" vs. "Sibling", χ^2 (1, N = 2648) = 0.0191571, p = 0.9118 (which is much greater than alpha = 0.05). Therefore, both teacher and sibling were equally number one helper for SIM lessons.

Effect Size

The test statistic is Q = 1670.831, our sample size for SIM students is 2648 and we have five options (variables) for helpers of SIM lessons. Therefore, the effect size for this can be calculated by eta-squared (η^2) (Serlin, Carr, & Marascuilo, 1982). η^2 = 1670.831/((5-1)x2648) = 0.16, which is a large effect size.

Evidence on SIM Students' Help Sought for SIM Lessons

At least 89% of SIM students sought help for SIM lessons as there is statistically significant evidence (p = 0.035444) that the percentage of SIM students who sought help for SIM lessons is greater than population hypothesized value of 89%. In other words, a binomial test indicated that the percentage of SIM students who sought help for SIM lessons (N_{help} = 2386, 90.1%) was statistically significantly greater than the population hypothesized value of 89%, p = 0.035444 (which is less than significance level alpha = 0.05).

Also, Cochran's Q test indicated that there are differences between the proportions among the five options of help for SIM lessons, $\chi^2(4, N = 2648) = 1670.831$, p =0.0000, with a large effect size (η^2 = 0.16). An exact pairwise post-hoc Cochran's Q test was not statistically significant for "Teacher" vs. "Sibling", $\chi^2 (1, N = 2648) = 0.0191571$, p = 0.9118 (which is much greater than alpha = 0.05). Therefore, both teacher and sibling were equally number one helper for SIM lessons.

Comparison between SIM Learning and Classroom Learning

Effectiveness of SIM learning vs Classroom Learning in increasing knowledge

The Ministry of Education was interested to know how SIM students found SIM learning in increasing their knowledge. To investigate this, ***Figure 22*** shows the results of SIM students' perception on increasing their knowledge during SIM learning in comparison to classroom learning.

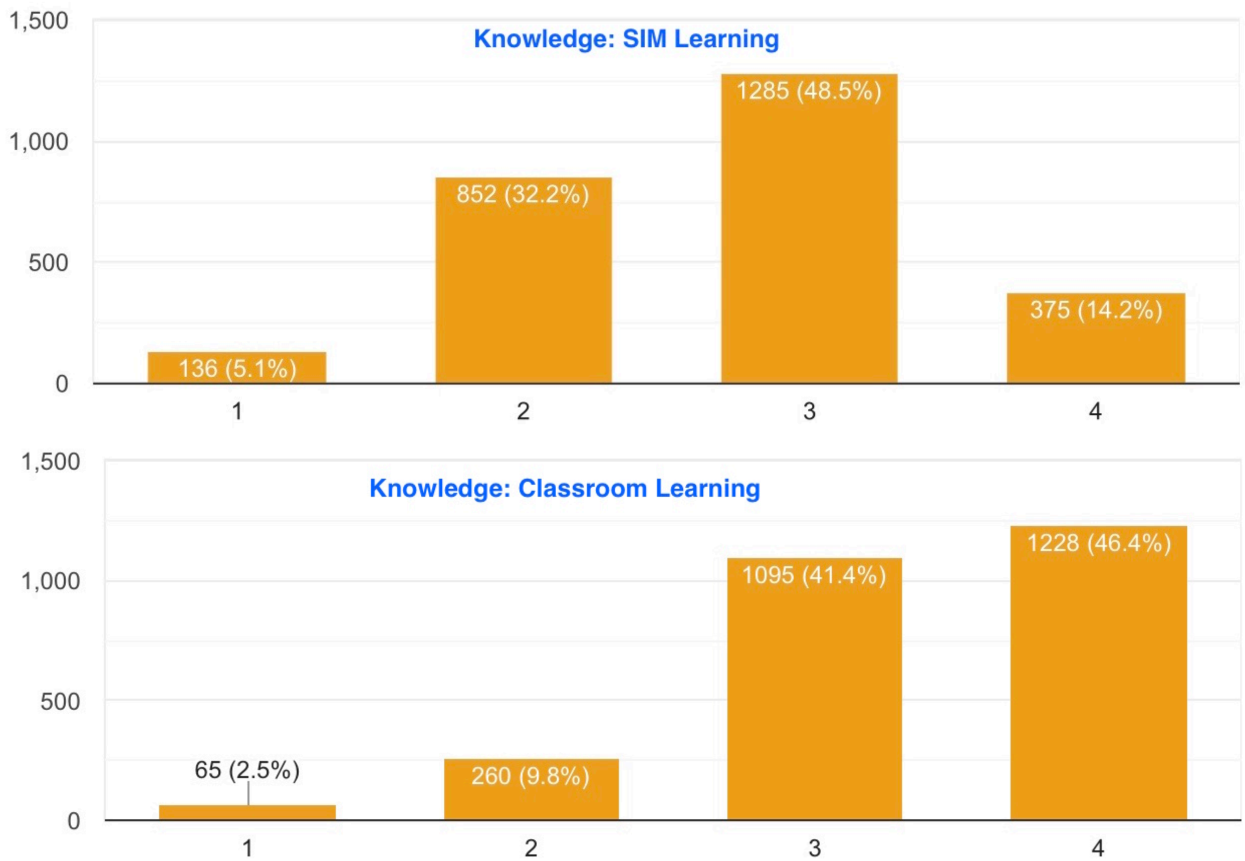


Figure 22: *Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of increasing knowledge" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 22*** the 62.7% (SIM) vs 87.8% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in increasing their knowledge.

Inferential Analysis – Statistical Significance Testing

Table 129: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q7 = q14

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      232      409543   1508860.5
    Negative |     1427     2608178   1508860.5
        Zero |      989      489555      489555
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -39247970
Adjustment for zeros  -80735779
                     ----------
Adjusted variance     1.428e+09

H0: q7 = q14
         z = -29.089
Prob > |z| =  0.0000
```

We have seen that the 62.7% of SIM students surveyed think that SIM learning was effective or extremely effective in increasing their knowledge. Comparing it with classroom learning, 87.8% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in increasing knowledge. Classroom learning was more effective in increasing knowledge. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for increasing knowledge is significantly different from the true median of classroom learning in increasing knowledge in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of increasing knowledge.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of increasing knowledge.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of increasing knowledge. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -29.089, p = 0.0000.

The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

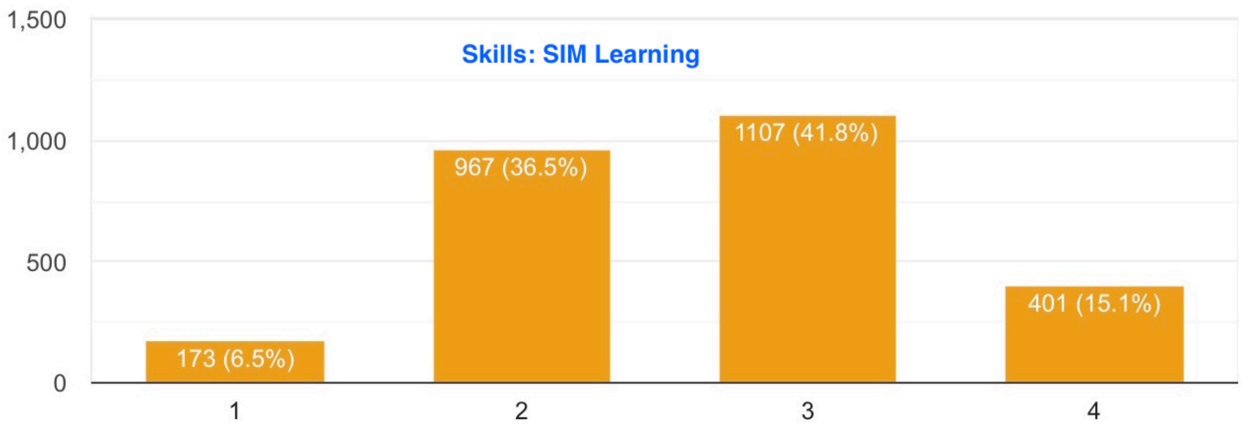
The test statistic is Z = -29.089 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -29.089/SQRT(2648) = -0.57 or 0.57 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Increasing Knowledge

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in increasing knowledge. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of increasing knowledge, Z = -29.089, p = 0.0000, with a moderate effect size or difference (*r* = 0.57).

Effectiveness of SIM learning vs Classroom Learning in increasing skills

The Ministry of Education was interested to know how SIM students found SIM learning in increasing their skills. To investigate this, ***Figure 23*** shows the results of SIM students' perception on increasing their skills during SIM learning in comparison to classroom learning.



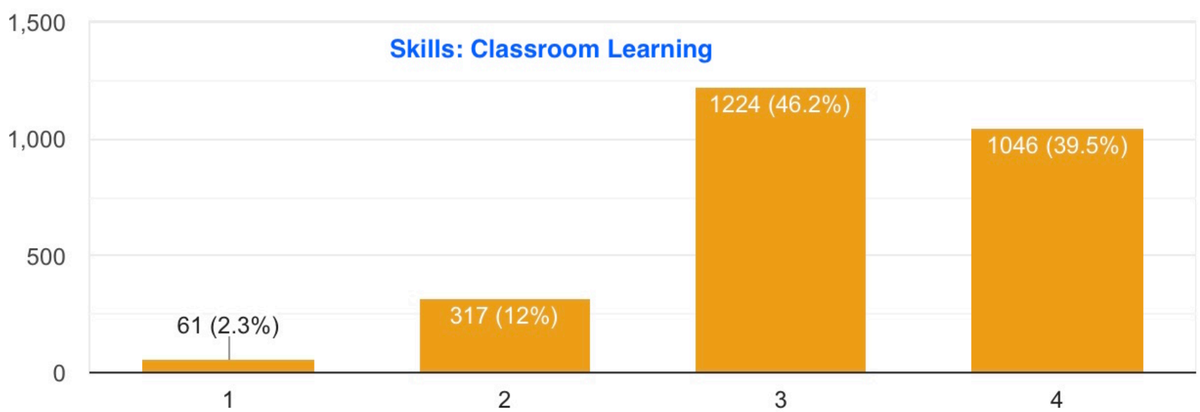


Figure 23*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of increasing skills" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 23*** the 56.9% (SIM) vs 85.7% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in increasing their skills.

Inferential Analysis – Statistical Significance Testing

Table 130: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = q15

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      288    495823.5     1515250
    Negative |     1384   2534676.5     1515250
        Zero |      976      476776      476776
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -38585827
Adjustment for zeros  -77595294
                     ----------
Adjusted variance     1.432e+09

H0: q8 = q15
         z = -26.939
Prob > |z| =  0.0000
```

We have seen that the 56.9% of SIM students surveyed think that SIM learning was effective or extremely effective in increasing their skills. Comparing it with classroom learning, 85.7% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in increasing skills. Classroom learning was more effective in increasing skills. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for increasing skills is significantly different from the true median of classroom learning in increasing skills in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of increasing skills.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of increasing skills.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of increasing skills. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -26.939, p = 0.0000. The negative

z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -26.939 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -26.939/SQRT(2648) = -0.52 or 0.52 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Increasing Skills

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in increasing skills. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of increasing skills, Z = -26.939, p = 0.0000, with a moderate effect size or difference ($r = 0.52$).

Effectiveness of SIM learning vs Classroom Learning in imparting values

The Ministry of Education was interested to know how SIM students found SIM learning in imparting values. To investigate this, ***Figure 24*** shows the results of SIM students' perception on imparting values during SIM learning in comparison to classroom learning.

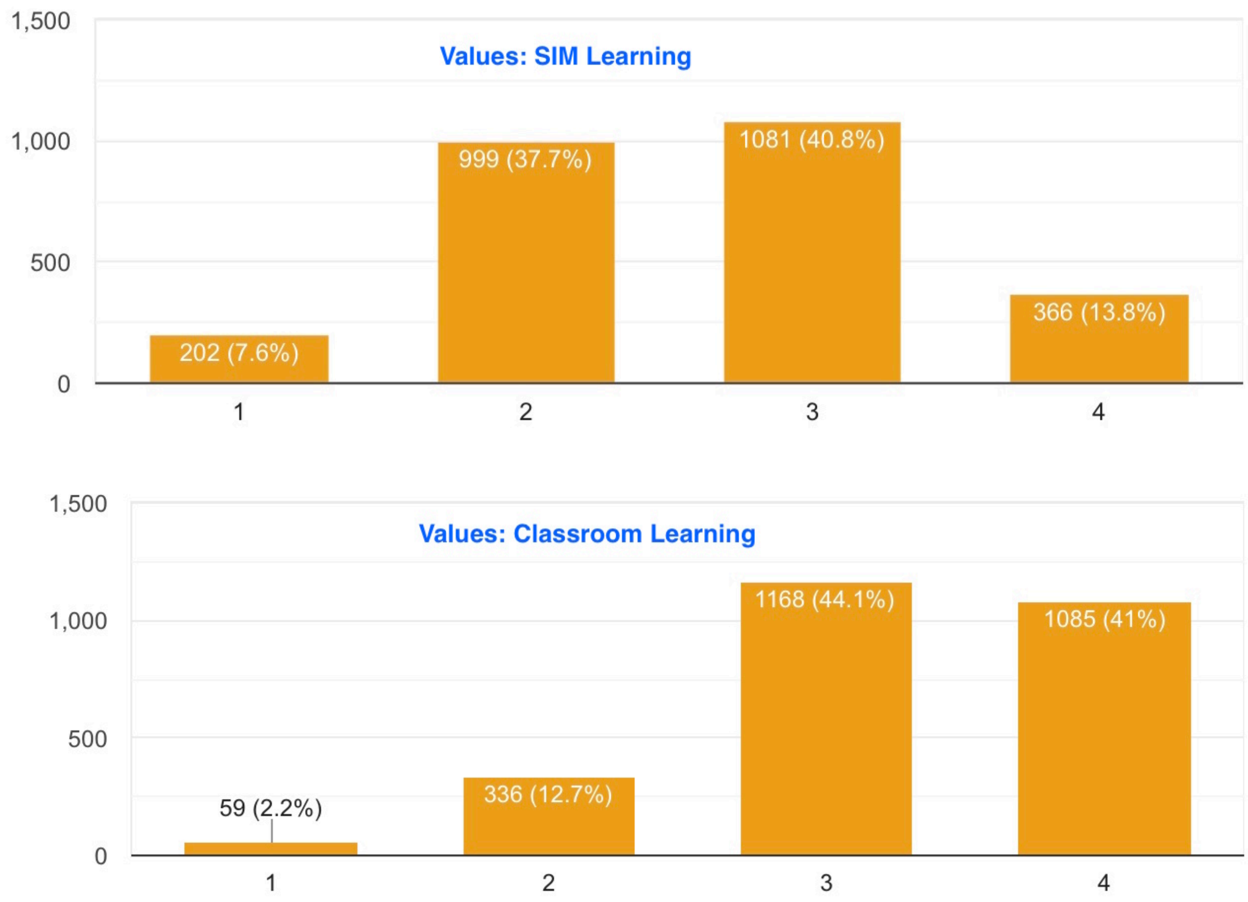


Figure 24*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of imparting values" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 24*** the 54.6% (SIM) vs 85.1% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in imparting values.

Inferential Analysis – Statistical Significance Testing

Table 131: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q9 = q16

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      270      450565     1528725
    Negative |     1430     2606885     1528725
        Zero |      948      449826      449826
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -35498216
Adjustment for zeros  -71109994
                     ----------
Adjusted variance     1.442e+09

H0: q9 = q16
         z = -28.397
Prob > |z| =  0.0000
```

We have seen that the 54.6% of SIM students surveyed think that SIM learning was effective or extremely effective in imparting values. Comparing it with classroom learning, 85.1% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in imparting values. Classroom learning was more effective in imparting values. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for imparting values is significantly different from the true median of classroom learning in imparting values in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of imparting values.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of imparting values.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of imparting values. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -28.397, p = 0.0000. The negative

z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -28.397 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -28.397/SQRT(2648) = -0.55 or 0.55 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Imparting Values

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in imparting values. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of imparting values, Z = -28.397, p = 0.0000, with a moderate effect size or difference (*r* = 0.55).

Effectiveness of SIM learning vs Classroom Learning in improving attitudes

The Ministry of Education was interested to know how SIM students found SIM learning in improving attitudes. To investigate this, ***Figure 25*** shows the results of SIM students' perception on improving attitudes during SIM learning in comparison to classroom learning.

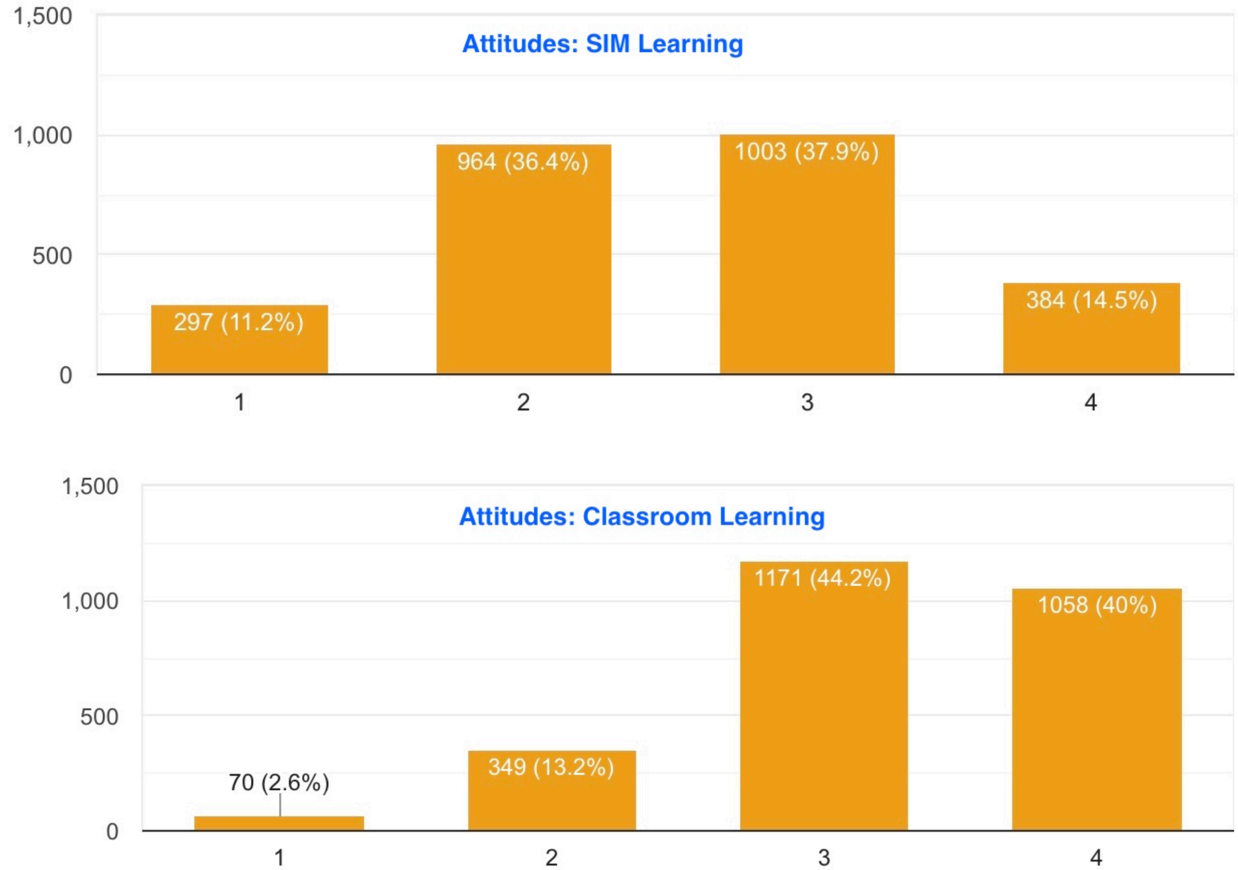


Figure 25: *Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of improving attitudes" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 25*** the 52.4% (SIM) vs 84.2% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in improving attitudes.

Inferential Analysis – Statistical Significance Testing

Table 132: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q10 = q17

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      282      464875   1533911.5
    Negative |     1429     2602948   1533911.5
        Zero |      937      439453      439453
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -32699720
Adjustment for zeros  -68664531
                     ----------
Adjusted variance     1.447e+09

H0: q10 = q17
         z = -28.105
Prob > |z| =  0.0000
```

We have seen that the 52.4% of SIM students surveyed think that SIM learning was effective or extremely effective in improving attitudes. Comparing it with classroom learning, 84.2% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in improving attitudes. Classroom learning was more effective in improving attitudes. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for improving attitudes is significantly different from the true median of classroom learning in improving attitudes in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of improving attitudes.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of improving attitudes.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of improving attitudes. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -28.105, p = 0.0000. The negative

z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

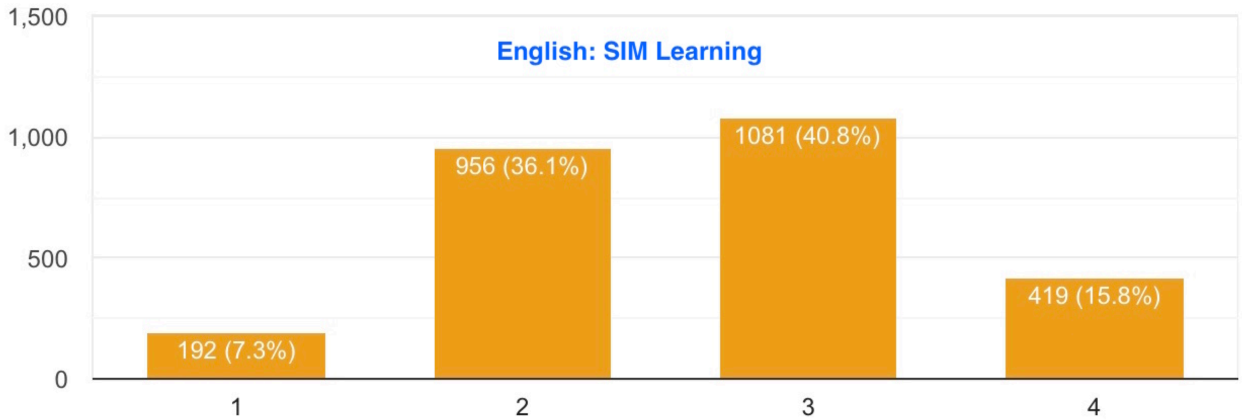
The test statistic is Z = -28.105 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -28.105/SQRT(2648) = -0.55 or 0.55 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Improving Attitudes

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in improving attitudes. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of improving attitudes, Z = -28.105, p = 0.0000, with a moderate effect size or difference (*r* = 0.55).

Effectiveness of SIM learning vs Classroom Learning in understanding English

The Ministry of Education was interested to know how SIM students found SIM learning in understanding English. To investigate this, ***Figure 26*** shows the results of SIM students' perception on understanding English during SIM learning in comparison to classroom learning.



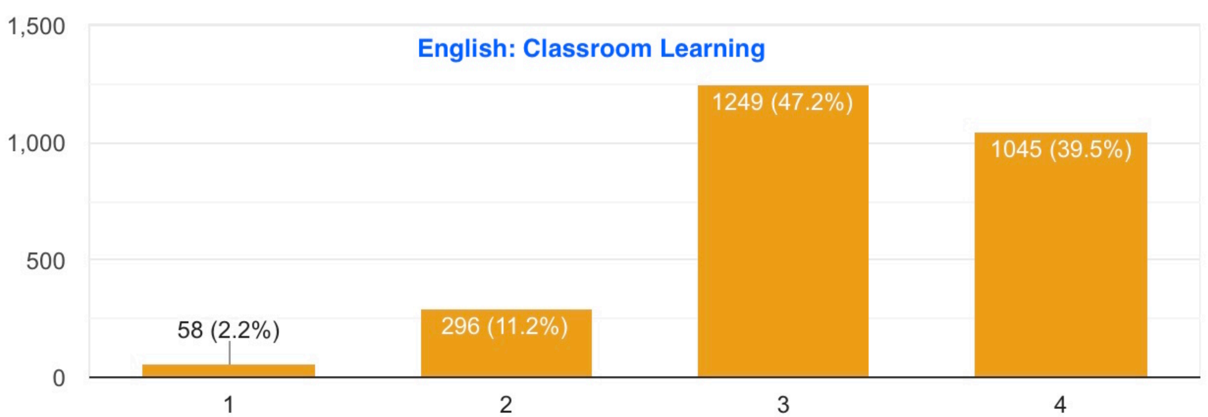


Figure 26*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of understanding English" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 26*** the 56.6% (SIM) vs 86.7% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in understanding English.

Inferential Analysis – Statistical Significance Testing

Table 133: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q11 = q18

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      221      390634     1480893
    Negative |     1383     2571152     1480893
        Zero |     1044      545490      545490
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -36065245
Adjustment for zeros  -94960718
                     ----------
Adjusted variance     1.417e+09

H0: q11 = q18
         z = -28.962
Prob > |z| =  0.0000
```

We have seen that the 56.6% of SIM students surveyed think that SIM learning was effective or extremely effective in understanding English. Comparing it with classroom learning, 86.7% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in understanding English. Classroom learning was more effective in understanding English. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for understanding English is significantly different from the true median of classroom learning in understanding English in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of understanding English.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of understanding English.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of understanding English. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = -28.962$, $p = 0.0000$.

The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -28.962 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -28.962/SQRT(2648) = -0.56 or 0.56 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Understanding English

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in understanding English. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of understanding English, Z = -28.962, p = 0.0000, with a moderate effect size or difference (*r* = 0.56).

Effectiveness of SIM learning vs Classroom Learning in understanding Maths

The Ministry of Education was interested to know how SIM students found SIM learning in understanding Mathematics. To investigate this, ***Figure 27*** shows the results of SIM students' perception on understanding Mathematics during SIM learning in comparison to classroom learning.

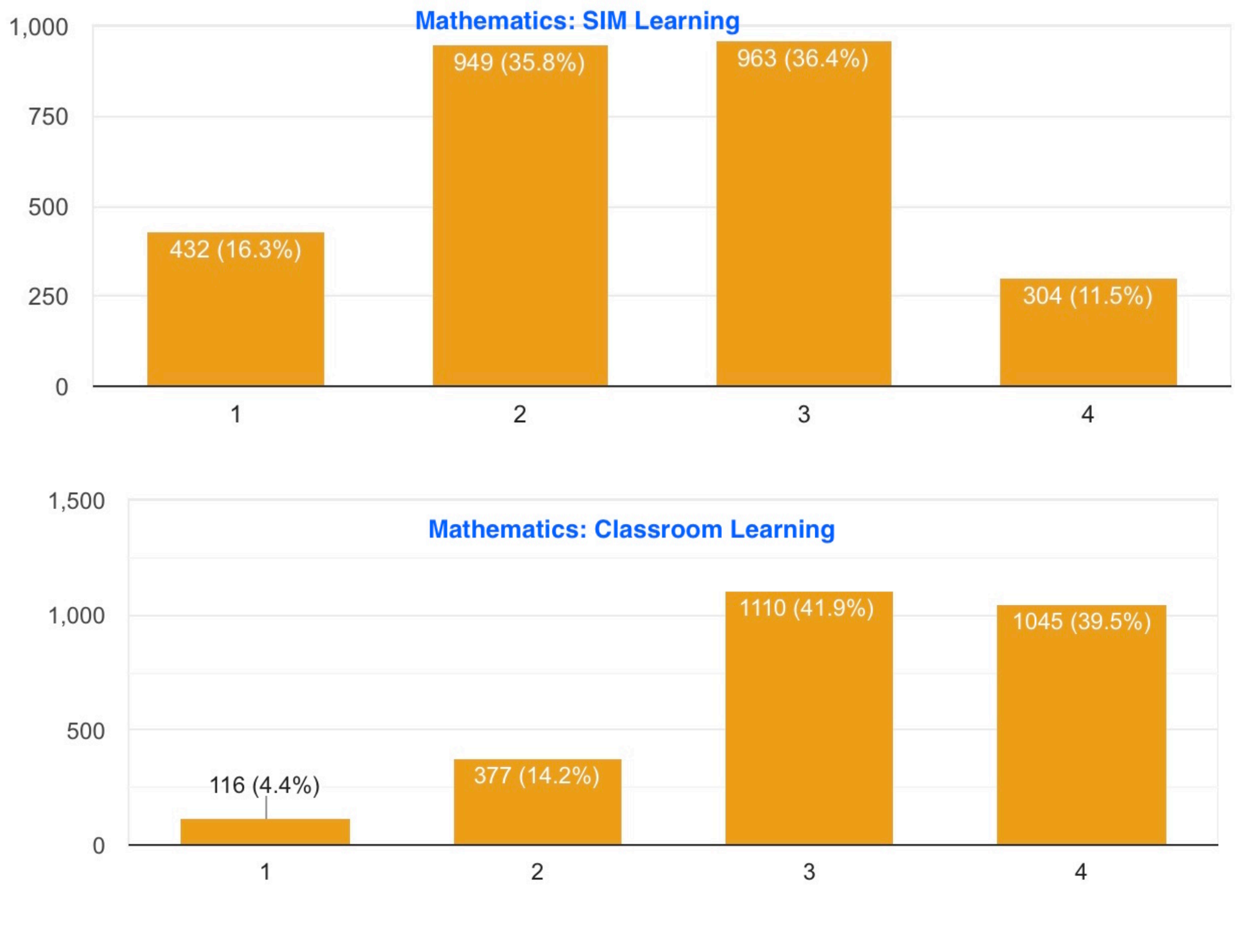


Figure 27*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of understanding Mathematics" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 27*** only 47.9% (SIM) vs 81.4% (Classroom) of the SIM student respondents rated "effective" or "extremely effective" in understanding Mathematics.

Inferential Analysis – Statistical Significance Testing

Table 134: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q12 = q19

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      188    316095.5   1504387.5
    Negative |     1462   2692679.5   1504387.5
        Zero |      998      498501      498501
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -25695450
Adjustment for zeros  -82958875
                     ----------
Adjusted variance     1.440e+09

H0: q12 = q19
         z = -31.320
Prob > |z| =  0.0000
```

We have seen that the 47.9% of SIM students surveyed think that SIM learning was effective or extremely effective in understanding Mathematics. Comparing it with classroom learning, 81.4% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in understanding Mathematics. Classroom learning was more effective in understanding Mathematics. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for understanding Mathematics is significantly different from the true median of classroom learning in understanding Mathematics in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of understanding Mathematics.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of understanding Mathematics.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of understanding Mathematics. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -31.320, p = 0.0000. The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -31.320 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -31.320/SQRT(2648) = -0.61 or 0.61 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Understanding Maths

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in understanding Mathematics. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of understanding Mathematics, Z = -31.320, p = 0.0000, with a strong effect size or difference (*r* = 0.61).

Effectiveness of SIM learning vs Classroom Learning in understanding Dzongkha

The Ministry of Education was interested to know how SIM students found SIM learning in understanding Dzongkha. To investigate this, ***Figure 28*** shows the results of SIM students' perception on understanding Dzongkha during SIM learning in comparison to classroom learning.

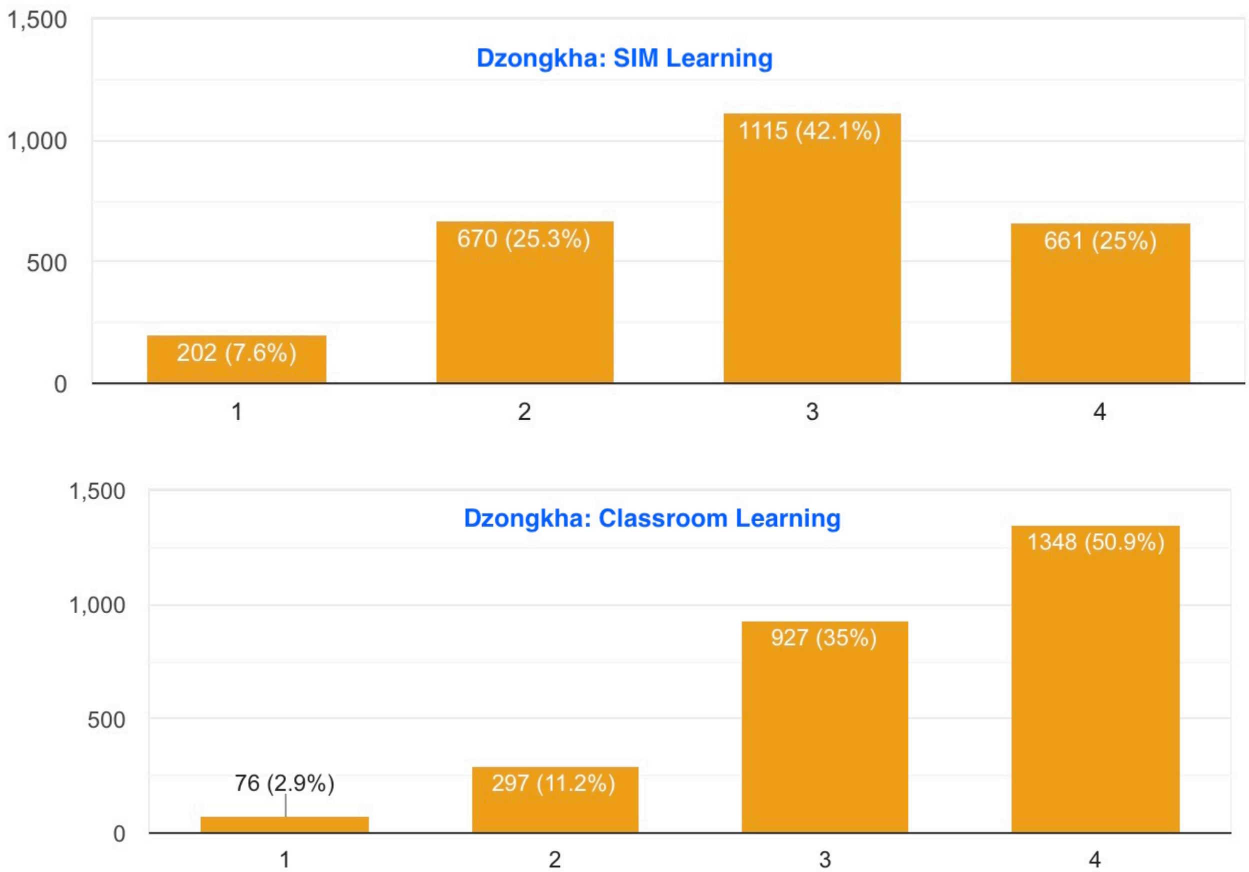


Figure 28*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of understanding Dzongkha" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 28*** the 67.1% (SIM) vs 85.9% (Classroom) majority of the SIM student respondents rated "effective" or "extremely effective" in understanding Dzongkha.

Inferential Analysis – Statistical Significance Testing

Table 135: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q13 = q20

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      221      417376     1398720
    Negative |     1236     2380064     1398720
        Zero |     1191      709836      709836
-------------+---------------------------------
         All |     2648     3507276     3507276

Unadjusted variance   1.548e+09
Adjustment for ties   -29259331
Adjustment for zeros -1.410e+08
                     ----------
Adjusted variance     1.378e+09

H0: q13 = q20
         z = -26.437
Prob > |z| =  0.0000
```

We have seen that the 67.1% of SIM students surveyed think that SIM learning was effective or extremely effective in understanding Dzongkha. Comparing it with classroom learning, 85.9% of the same group of SIM students surveyed also think that classroom learning was effective or extremely effective in understanding Dzongkha. Classroom learning was more effective in understanding Dzongkha. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM student population too. In other words, we have to test whether the true median of SIM learning for understanding Dzongkha is significantly different from the true median of classroom learning in understanding Dzongkha in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of understanding Dzongkha.

Ha: Our alternative hypothesis is that the SIM student population's true median of SIM learning is significantly different from true median of classroom learning in terms of understanding Dzongkha.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of understanding Dzongkha. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -26.437, p = 0.0000.

The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -26.437 and our sample size for SIM students is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -26.437/SQRT(2648) = -0.51 or 0.51 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size or difference.

Evidence on SIM Learning vs Classroom Learning in Understanding Dzongkha

There is statistically significant evidence (p = 0.0000) that the SIM students found classroom learning more effective than SIM learning in understanding Dzongkha. In particular, two-sample Wilcoxon signed rank test indicated that the students tend to like classroom learning more than SIM learning in terms of understanding Dzongkha, Z = -26.437, p = 0.0000, with a moderate effect size or difference (*r* = 0.51).

PART II: SIM TEACHERS

Demographic Characteristics of SIM Teacher Respondents

The age characteristics of the SIM teacher respondents are summarized in ***Table 136***. The age of the SIM teacher respondents ranged from 24 to 57 years (M = 33.85, SD = 6.45).

Table 136: Results of age characteristics of SIM teacher respondents

Variable	Obs	Mean	Std. dev.	Min	Max
age	667	33.85157	6.451518	24	57

Similarly, among the 667 SIM teacher respondents, 400 (60%) were males and 267 (40%) were females as shown in ***Figure 29***.

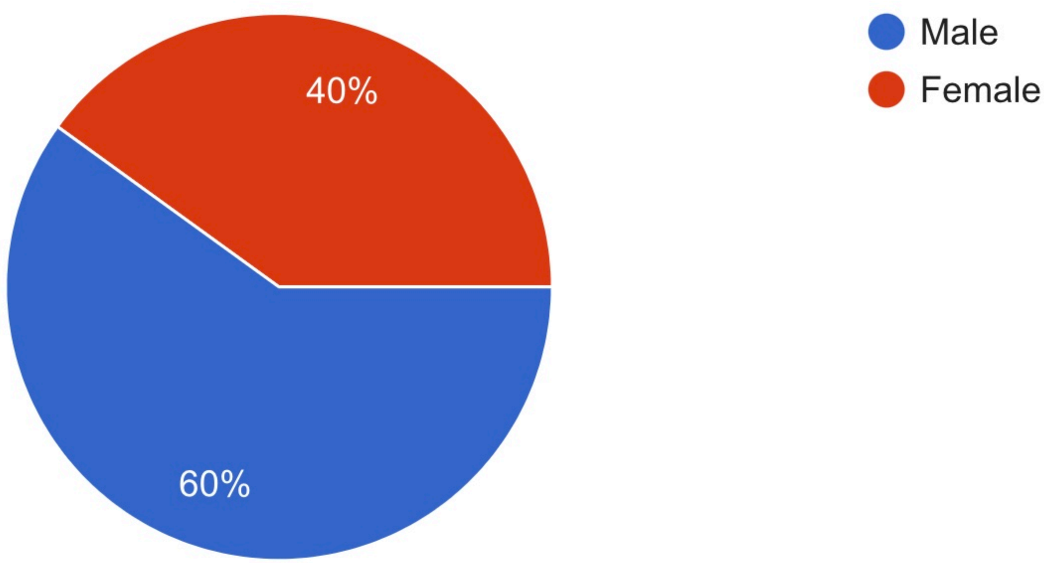


Figure 29: *Gender of SIM teacher respondents*

Likewise, among the 667 SIM teacher respondents, we got data representation from all classes from Class PP to Class XII as shown in ***Figure 30***, with maximum teaching class X (18.3%), followed by class XII (13.2%), class VI (12.1%), class I (7.8%), class III (6.5%), class IV (6.3%), class V (6.3%), class IX (6.3%), class VIII (6.0%), class VII (5.9%), class II (5.3%), class XI (6.1%) and minimum teaching class PP (1.2%).

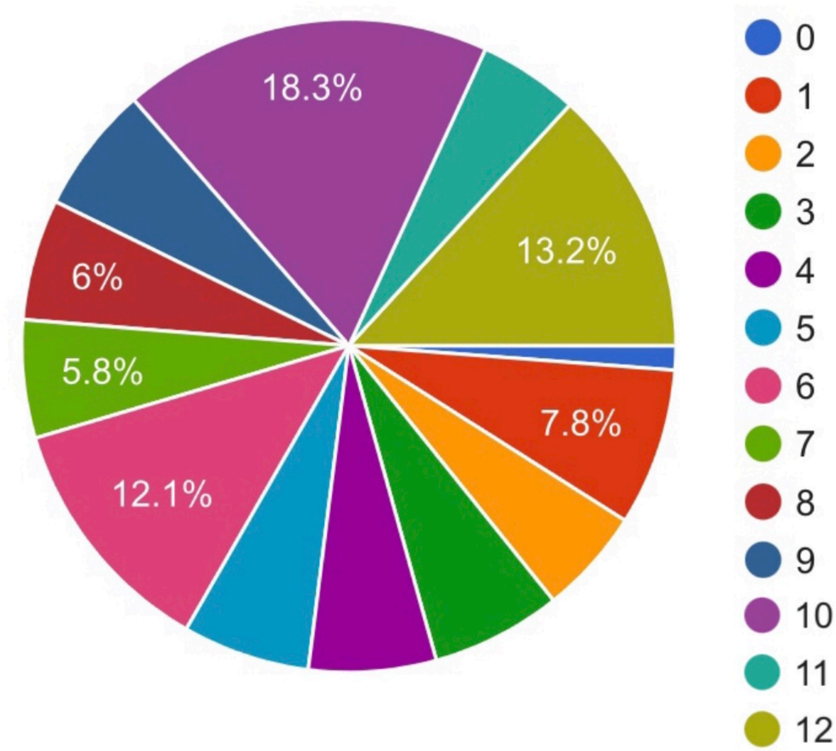


Figure 30*: Results of "What class do you teach mainly?"*

Among the 667 SIM teacher respondents, we got data representation from all types of schools such as HSS (43.5%), MSS (29.2%), LSS (7.8%), PS (18.7%), and ECR (0.8%) as shown in ***Figure 31***.

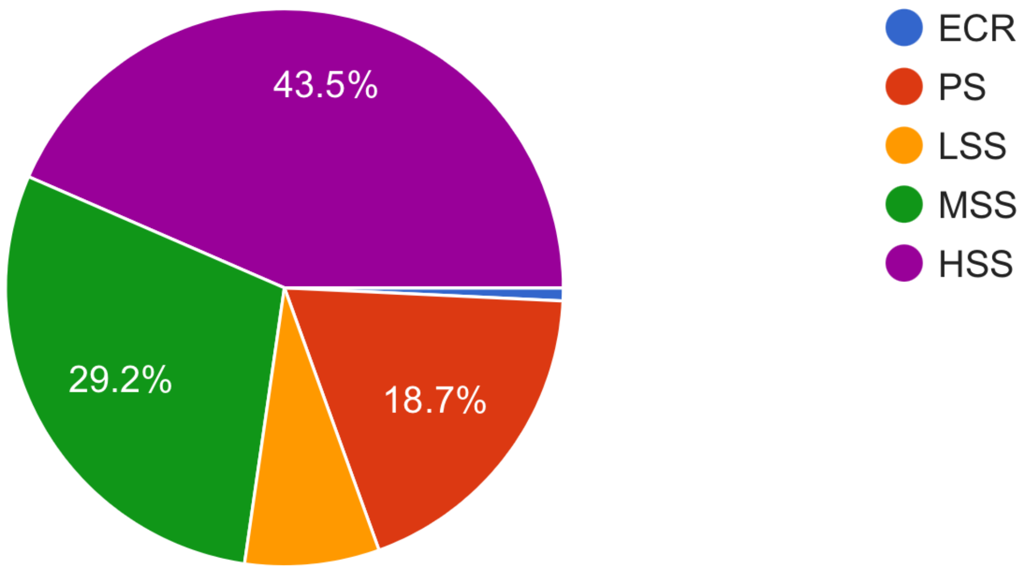


Figure 31*: School types of SIM teacher respondents*

Effectiveness of SIM Programme

Analyzing Teachers' Satisfaction Level of SIM

The Ministry of Education was interested to know satisfaction level of SIM programme, including SIM teachers' satisfaction level, during COVID-19 pandemic. To investigate this, ***Figure 32***, which is visualization of survey data, shows the results of satisfaction opinion from the SIM survey.

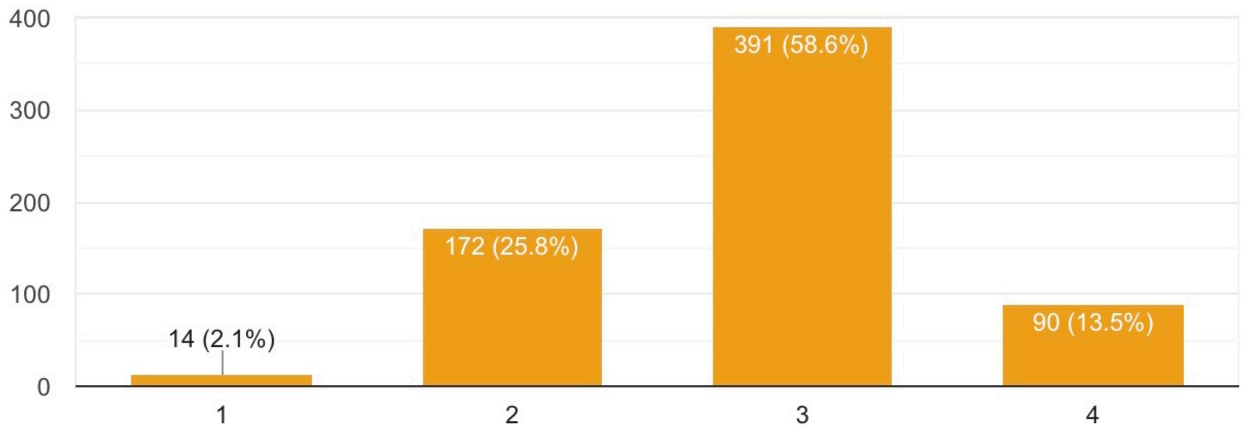


Figure 32: *Results of "Rate how satisfied are you with the current SIM" where 1 = Extremely dissatisfied, 2 = Dissatisfied, 3 = Satisfied, and 4 = Extremely satisfied*

As can be seen in ***Figure 32*** the 72.1% of the SIM teacher respondents rated the SIM programme "satisfied" or "extremely satisfied."

Descriptive Analysis – Measure of Central Tendency

Table 137: Results of the SIM teachers' satisfaction level rating frequency distribution

```
. tabulate q27

        q27 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         14        2.10        2.10
          2 |        172       25.79       27.89
          3 |        391       58.62       86.51
          4 |         90       13.49      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency ***Table 137*** above, it shows that mode choice is 3, which is "satisfied." The total SIM teacher respondents of 72.1% chose "satisfied" or "extremely satisfied."

Table 138: SIM teachers' satisfaction level rating frequency distribution, by age group

```
. tabulate age_group q27

           |                     q27
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         0          1          3          1 |         5
   (25-29) |         6         55        113         14 |       188
   (30-34) |         5         53        120         25 |       203
   (35-39) |         2         40         81         31 |       154
   (40-44) |         1         13         40         11 |        65
   (45-49) |         0          7         22          7 |        36
   (50-54) |         0          3          9          0 |        12
   (55-59) |         0          0          3          1 |         4
-----------+--------------------------------------------+----------
     Total |        14        172        391         90 |       667
```

Looking at teachers' satisfaction level of SIM survey data by age group, it shows that consistently in all age groups, the mode or most choice selected is 3, which is "satisfied."

Table 139: SIM teachers' satisfaction level rating frequency distribution, by key stage

```
. tabulate key_stage q27

              |                     q27
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         0         34         83         21 |       138
 Key Stage II |         2         27        110         26 |       165
Key Stage III |         4         19         47          9 |        79
 Key Stage IV |         4         52         89         19 |       164
  Key Stage V |         4         40         62         15 |       121
--------------+--------------------------------------------+----------
        Total |        14        172        391         90 |       667
```

Similarly, looking at teachers' satisfaction level of SIM survey data by key stage, it shows that consistently in all key stages, the mode or most choice selected is 3, which is "satisfied."

Table 140: SIM teachers' satisfaction level rating frequency distribution, by school type

```
. tabulate school q27

           |                         q27
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          1          3          1 |         5
       HSS |         6         84        155         45 |       290
       LSS |         1          8         36          7 |        52
       MSS |         7         60        111         17 |       195
        PS |         0         19         86         20 |       125
-----------+--------------------------------------------+----------
     Total |        14        172        391         90 |       667
```

Likewise, looking at teachers' satisfaction level of SIM survey data by school type, it shows that consistently in all school types, the mode or most choice selected is 3, which is "satisfied."

Table 141: Result of the SIM teachers' satisfaction level rating median calculation

```
. tabstat q27, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q27 |       667         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "satisfied." This means at least 50% of the SIM teacher respondents are in the "satisfied" or "extremely satisfied" category looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 142: Result of the SIM teachers' measure of consensus on satisfaction level

```
. cns q27 , min(1) max(4)

Consensus Measure for q27
Cns(X) = .70648676
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the satisfaction level of SIM teachers, it is 0.7064.

Inferential Analysis – Statistical Significance Testing

Table 143: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q27 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      481      165657      111389
    Negative |      186       57121      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties    -3741203
Adjustment for zeros          0
                     ----------
Adjusted variance      21042850

H0: q27 = 2.5
         z = 11.830
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 72.1% of SIM teachers surveyed think that SIM programme was satisfactory. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "dissatisfied" and 3 = "satisfied."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 11.830, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 11.830 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 11.830/SQRT(667) = 0.46. This, according to Bartz (1999) is moderate effect size.

Gender difference in satisfaction level of SIM learning

Table 144: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q27, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267     91378.5       89178
        Male |      400    131399.5      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.31e+06
                     ----------
Adjusted variance    4630982.41

H0: q27(gender==Female) = q27(gender==Male)
         z =  1.023
Prob > |z| = 0.3065
Exact prob = 0.3076
```

There is no evidence for statistically significant difference between satisfaction level of SIM learning between female teachers and male teachers (p-value = 0.3076 > alpha = 0.05), which means both female teachers and male teachers are equally satisfied with SIM learning.

Evidence on SIM Teachers' Satisfaction Level

There is statistically significant evidence (p = 0.0000) that the majority 72.1% of SIM teachers, both female teachers and male teachers, are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 11.830, p = 0.0000, with a moderate effect size (*r* = 0.46).

Analyzing Teachers' Acceptance Level of SIM

The Ministry of Education was interested to know acceptance level of SIM programme, including SIM teachers' acceptance level, during COVID-19 pandemic. To investigate this, ***Figure 33*** shows the results of SIM acceptance opinion from the SIM survey.

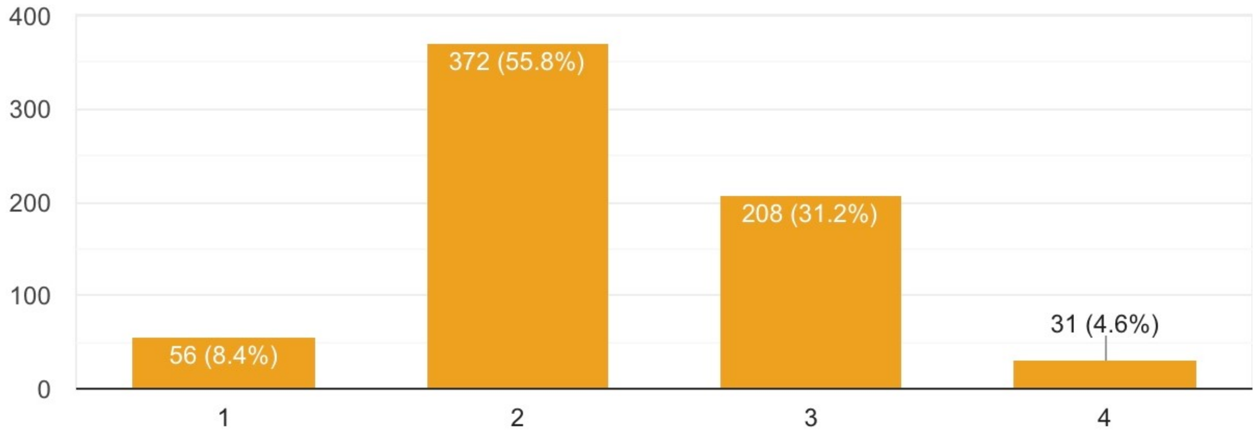


Figure 33*: Results of "Rate how much did your students enjoy SIM learning during the pandemic" where 1 = Extremely unenjoyable, 2 = Unenjoyable, 3 = Enjoyable, and 4 = Extremely enjoyable*

As can be seen in ***Figure 33*** only 35.8% of the SIM teacher respondents rated the SIM learning "enjoyable" or "extremely enjoyable" for their students.

Descriptive Analysis – Measure of Central Tendency

Table 145: Results of the SIM teachers' acceptance level rating frequency distribution

```
. tabulate q21

        q21 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         56        8.40        8.40
          2 |        372       55.77       64.17
          3 |        208       31.18       95.35
          4 |         31        4.65      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "unenjoyable." The SIM teacher respondents of only 35.8% chose SIM "enjoyable" or "extremely enjoyable" for their students.

Table 146: SIM teachers' acceptance level rating frequency distribution, by age group

```
. tabulate age_group q21

           |                         q21
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         1          3          1          0 |         5
   (25-29) |        17        112         54          5 |       188
   (30-34) |        17        111         67          8 |       203
   (35-39) |        12         81         49         12 |       154
   (40-44) |         4         39         18          4 |        65
   (45-49) |         4         19         12          1 |        36
   (50-54) |         1          5          5          1 |        12
   (55-59) |         0          2          2          0 |         4
-----------+--------------------------------------------+----------
     Total |        56        372        208         31 |       667
```

Looking at teachers' acceptance level of SIM survey data by age group, it shows that consistently almost in all age groups, the mode or most choice selected is 2, which is "unenjoyable."

Table 147: SIM teachers' acceptance level rating frequency distribution, by key stage

```
. tabulate key_stage q21

              |                         q21
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         6         68         56          8 |       138
 Key Stage II |        11         80         59         15 |       165
Key Stage III |         9         50         17          3 |        79
 Key Stage IV |        19         98         46          1 |       164
  Key Stage V |        11         76         30          4 |       121
--------------+--------------------------------------------+----------
        Total |        56        372        208         31 |       667
```

Similarly, looking at teachers' acceptance level of SIM survey data by key stage, it shows that consistently in all key stages, the mode or most choice selected is 2, which is "unenjoyable."

Table 148: SIM teachers' acceptance level rating frequency distribution, by school type

```
. tabulate school q21

           |                       q21
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          0          4          1 |         5
       HSS |        27        178         76          9 |       290
       LSS |         3         30         16          3 |        52
       MSS |        24        112         52          7 |       195
        PS |         2         52         60         11 |       125
-----------+--------------------------------------------+----------
     Total |        56        372        208         31 |       667
```

Likewise, looking at teachers' acceptance level of SIM survey data by school type, it shows that in higher level school types such as HSS, MSS and LSS, the mode or most choice selected is 2, which is "unenjoyable" but teachers of lower level school types such as PS and ECR have selected most choice as 3 which is "enjoyable." So it was a mixed response.

Table 149: Result of the SIM teachers' acceptance level rating median calculation

```
. tabstat q21, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q21 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "unenjoyable." This means at least 50% of the SIM teacher respondents are in the "unenjoyable" or "extremely unenjoyable" group looking at the median score rating of 2.

Descriptive Analysis - Measure of Dispersion

Table 150: Result of the SIM teachers' measure of consensus on acceptance level

```
. cns q21 , min(1) max(4)

Consensus Measure for q21
Cns(X) = .66821235
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the acceptance level of SIM teachers, it is 0.6682.

Inferential Analysis – Statistical Significance Testing

Table 151: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q21 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      239       79768      111389
    Negative |      428      143010      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -4078538.3
Adjustment for zeros          0
                     ----------
Adjusted variance      20705514

H0: q21 = 2.5
         z = -6.949
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that only 35.8% of SIM teachers surveyed think that SIM programme was enjoyable or extremely enjoyable. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = “unenjoyable” and 3 = “enjoyable.”

Ho: Our null hypothesis is that the SIM teacher population’s true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population’s true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -6.949, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -6.949 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -6.949/SQRT(667) = - 0.27. Ignoring negative sign, this, according to Bartz (1999), is low effect size or difference.

Gender difference in acceptance level of SIM learning

Table 152: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q21, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267     84169.5       89178
        Male |      400    138608.5      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.22e+06
                     ----------
Adjusted variance    4729424.44

H0: q21(gender==Female) = q21(gender==Male)
         z = -2.303
Prob > |z| = 0.0213
Exact prob = 0.0211
```

There is evidence for statistically significant difference between acceptance level of SIM learning between female teachers and male teachers (p-value = 0.0211 < alpha = 0.05), which means female teachers and male teachers rated SIM learning acceptance level for their students differently.

Evidence on SIM Teachers' Acceptance Level

There is statistically significant evidence (p = 0.0000) that only 35.8% of SIM teachers found SIM learning enjoyable during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -6.949, p = 0.0000, with a low effect size (*r* = 0.27).

Effectiveness of SIM Materials

Analyzing Teachers' Perception on Overall Presentation of SIM Booklets

The Ministry of Education was interested to know how SIM teachers found overall presentation of the SIM booklets. To investigate this, ***Figure 34*** shows the results of SIM teachers' perception on overall presentation of the SIM booklets.

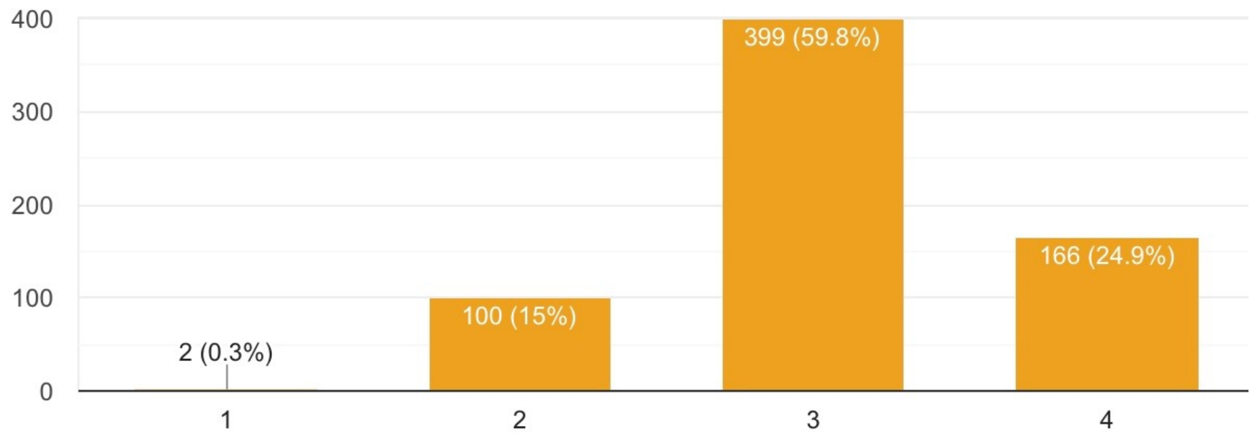


Figure 34*: Results of "Rate how did you find overall presentation of the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 34*** the 84.7% of the SIM teacher respondents rated the overall presentation of SIM materials "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 153: Results of the SIM teachers' rating of overall presentation of SIM frequency distribution

```
. tabulate q26

        q26 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          2        0.30        0.30
          2 |        100       14.99       15.29
          3 |        399       59.82       75.11
          4 |        166       24.89      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode is 3, which is "effective." The total SIM teacher respondents of 84.7% chose "effective" or "extremely effective."

Table 154: SIM teachers' rating of SIM overall presentation frequency distribution, by age group

```
. tabulate age_group q26
```

Age_Group	q26 1	2	3	4	Total
(20-24)	0	0	4	1	5
(25-29)	1	36	111	40	188
(30-34)	0	36	121	46	203
(35-39)	0	16	89	49	154
(40-44)	1	6	40	18	65
(45-49)	0	4	22	10	36
(50-54)	0	2	9	1	12
(55-59)	0	0	3	1	4
Total	2	100	399	166	667

Looking at teachers' rating of SIM overall presentation by age group, it shows consistently that all age groups have mode 3, which is "effective."

Table 155: SIM teachers' rating of SIM overall presentation frequency distribution, by key stage

```
. tabulate key_stage q26
```

Key_Stage	q26 1	2	3	4	Total
Key Stage I	0	19	77	42	138
Key Stage II	0	12	107	46	165
Key Stage III	0	15	44	20	79
Key Stage IV	1	26	100	37	164
Key Stage V	1	28	71	21	121
Total	2	100	399	166	667

Similarly, looking at teachers' rating of SIM overall presentation by key stage, it shows consistently that all key stages have mode 3, which is "effective."

Table 156: SIM teachers' rating of SIM overall presentation frequency distribution, by school type

```
. tabulate school q26

           |                    q26
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          1          3          1 |         5
       HSS |         1         49        173         67 |       290
       LSS |         0          6         29         17 |        52
       MSS |         1         35        114         45 |       195
        PS |         0          9         80         36 |       125
-----------+--------------------------------------------+----------
     Total |         2        100        399        166 |       667
```

Likewise, looking at teachers' rating of SIM overall presentation by school type, it shows that all school types rated SIM overall presentation as "effective."

Table 157: Result of the SIM teachers' rating of SIM overall presentation median calculation

```
. tabstat q26, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q26 |       667         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM teacher respondents found SIM overall presentation "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 158: Result of the SIM teachers' measure of consensus on SIM overall presentation rating

```
. cns q26 , min(1) max(4)

Consensus Measure for q26
Cns(X) = .7403962
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM overall presentation rating of SIM teachers, it is 0.7404.

Inferential Analysis – Statistical Significance Testing

Table 159: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q26 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      565      196611      111389
    Negative |      102       26167      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties    -2687343
Adjustment for zeros          0
                     ----------
Adjusted variance      22096710

H0: q26 = 2.5
         z = 18.130
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 84.7% of SIM teachers surveyed think that SIM overall presentation was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 18.130, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 18.130 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 18.130/SQRT(2648) = 0.70. This, according to Bartz (1999) is strong effect size.

Gender difference in SIM teachers' rating of SIM overall presentation

Table 160: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q26, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       93122       89178
        Male |      400      129656      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.38e+06
                     ----------
Adjusted variance    4560883.06

H0: q26(gender==Female) = q26(gender==Male)
         z =  1.847
Prob > |z| = 0.0648
Exact prob = 0.0651
```

There is no evidence for statistically significant difference in SIM overall presentation rating between female teachers and male teachers (p-value = 0.0651> alpha = 0.05), which means both female teachers and male teachers found SIM overall presentation equally effective.

Evidence on SIM Teachers' Perception of SIM Overall Presentation

There is statistically significant evidence (p = 0.0000) that the majority 84.7% of SIM teachers, both female teachers and male teachers, found overall presentation of the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 18.130, p = 0.0000, with a strong effect size (*r* = 0.70).

Analyzing Teachers' Perception on Contents in SIM Booklets

The Ministry of Education was interested to know how SIM teachers found contents of the SIM booklets. To investigate this, ***Figure 35*** shows the results of SIM teachers' perception on contents of the SIM booklets.

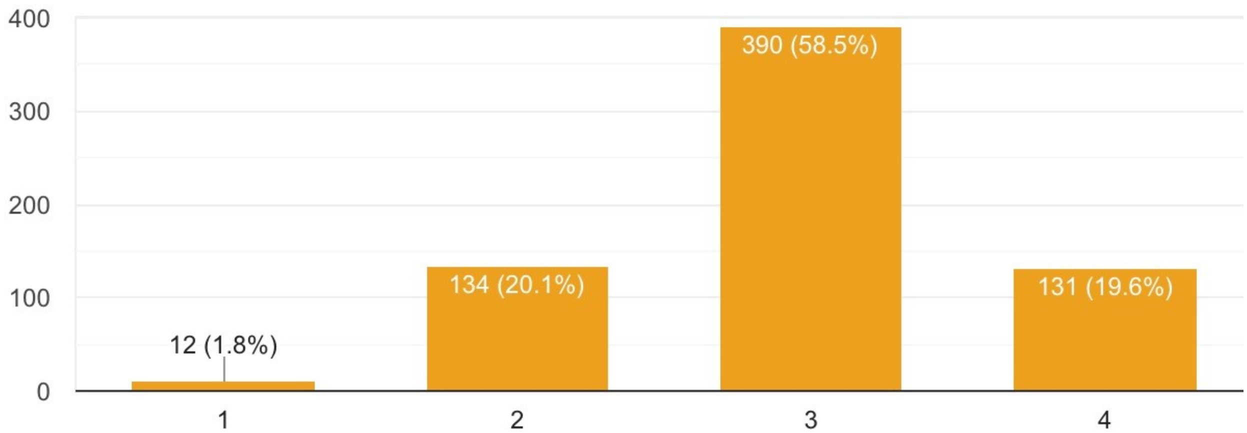


Figure 35*: Results of "Rate how did you find contents of the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 35*** the 78.1% of the SIM teacher respondents rated the SIM contents "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 161: Results of the SIM teachers' rating of SIM contents frequency distribution

```
. tabulate q23

        q23 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         12        1.80        1.80
          2 |        134       20.09       21.89
          3 |        390       58.47       80.36
          4 |        131       19.64      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM teacher respondents of 78.1% chose "effective" or "extremely effective."

Table 162: SIM teachers' rating of SIM contents frequency distribution, by age group

```
. tabulate age_group q23

           |                         q23
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         0          1          3          1 |         5
   (25-29) |         3         39        123         23 |       188
   (30-34) |         7         40        115         41 |       203
   (35-39) |         0         32         80         42 |       154
   (40-44) |         1         14         39         11 |        65
   (45-49) |         0          6         22          8 |        36
   (50-54) |         0          2          6          4 |        12
   (55-59) |         1          0          2          1 |         4
-----------+--------------------------------------------+----------
     Total |        12        134        390        131 |       667
```

Looking at teachers' rating of SIM contents by age group, it shows that in all age groups the mode choice selected is 3, which is "effective."

Table 163: SIM teachers' rating of SIM contents frequency distribution, by key stage

```
. tabulate key_stage q23

              |                         q23
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         4         28         78         28 |       138
 Key Stage II |         2         24        101         38 |       165
Key Stage III |         1         12         46         20 |        79
 Key Stage IV |         3         29        108         24 |       164
  Key Stage V |         2         41         57         21 |       121
--------------+--------------------------------------------+----------
        Total |        12        134        390        131 |       667
```

Similarly, looking at teachers' rating of SIM contents by key stage, it shows that consistently in all key stages, the mode is 3, which is "effective."

Table 164: SIM teachers' rating of SIM contents frequency distribution, by school type

```
. tabulate school q23

           |                         q23
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          1          3          1 |         5
       HSS |         5         64        166         55 |       290
       LSS |         0          8         28         16 |        52
       MSS |         7         37        119         32 |       195
        PS |         0         24         74         27 |       125
-----------+--------------------------------------------+----------
     Total |        12        134        390        131 |       667
```

Likewise, looking at teachers' rating of SIM contents by school type, it shows that consistently in all school types, the mode is 3, which is "effective."

Table 165: Result of the SIM teachers' rating of SIM contents median calculation

```
. tabstat q23, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q23 |       667         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM teacher respondents found SIM contents "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 166: Result of the SIM teachers' measure of consensus on SIM contents rating

```
. cns q23 , min(1) max(4)

Consensus Measure for q23
Cns(X) = .72866264
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM contents rating of SIM teachers, it is 0.7287.

Inferential Analysis – Statistical Significance Testing

Table 167: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q23 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      521      180451      111389
    Negative |      146       42327      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -3058361.8
Adjustment for zeros          0
                     ----------
Adjusted variance      21725691

H0: q23 = 2.5
         z = 14.817
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 78.1% of SIM teachers surveyed think that SIM contents was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.817 p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 14.817 and our sample size for SIM teachers is 2648. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 14.817/SQRT(667) = 0.57. This, according to Bartz (1999), is moderate effect size.

Gender difference in SIM teachers' rating of SIM contents

Table 168: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q23, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

       gender |      Obs    Rank sum    Expected
-------------+---------------------------------
       Female |      267       90928       89178
         Male |      400      131850      133600
-------------+---------------------------------
     Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.28e+06
                     ----------
Adjusted variance    4663474.02

H0: q23(gender==Female) = q23(gender==Male)
         z =  0.810
Prob > |z| = 0.4177
Exact prob = 0.4152
```

There is no evidence for statistically significant difference between SIM contents rating between female teachers and male teachers (p-value = 0.4152 > alpha = 0.05), which means both female teachers and male teachers found SIM contents equally effective.

Evidence on SIM Teachers' Perception of SIM Contents

There is statistically significant evidence (p = 0.0000) that the majority 78.1% of SIM teachers, both female teachers and male teachers, found contents of SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.817, p = 0.0000, with a moderate effect size (r = 0.57).

Analyzing Teachers' Perception on Instructions in SIM Booklets

The Ministry of Education was interested to know how SIM teachers found instructions incorporated in the SIM booklets. To investigate this, ***Figure 36*** shows the results of SIM teachers' perception on instructions in the SIM booklets.

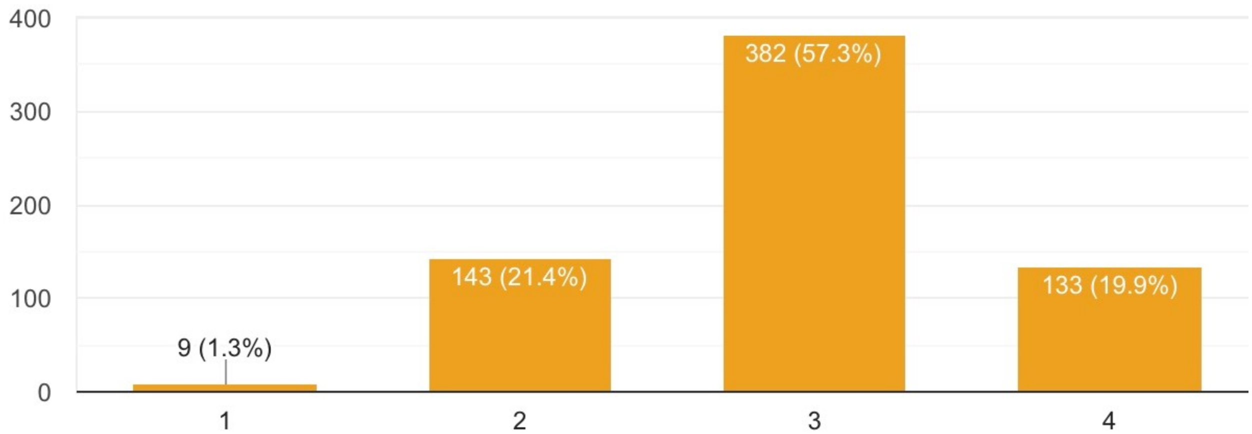


***Figure 36**: Results of "Rate how did you find instructions in the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 36*** the 77.2% of the SIM teacher respondents rated the SIM instructions "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 169: Results of the SIM teachers' rating of SIM instructions frequency distribution

```
. tabulate q22

        q22 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          9        1.35        1.35
          2 |        143       21.44       22.79
          3 |        382       57.27       80.06
          4 |        133       19.94      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM teacher respondents of 77.2% chose "effective" or "extremely effective."

Table 170: SIM teachers' rating of SIM instructions frequency distribution, by age group

```
. tabulate age_group q22

           |                         q22
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         0          1          4          0 |         5
   (25-29) |         3         50        104         31 |       188
   (30-34) |         5         46        116         36 |       203
   (35-39) |         1         26         84         43 |       154
   (40-44) |         0         11         40         14 |        65
   (45-49) |         0          7         24          5 |        36
   (50-54) |         0          2          9          1 |        12
   (55-59) |         0          0          1          3 |         4
-----------+--------------------------------------------+----------
     Total |         9        143        382        133 |       667
```

Looking at teachers' rating of SIM instructions by age group, it shows that in almost all age groups except age group 55-59, the mode or most choice selected is 3, which is "effective." The age group 55-59 has mode as 4, which is extremely effective.

Table 171: SIM teachers' rating of SIM instructions frequency distribution, by key stage

```
. tabulate key_stage q22

              |                         q22
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         0         27         79         32 |       138
 Key Stage II |         2         31         87         45 |       165
Key Stage III |         2         18         44         15 |        79
 Key Stage IV |         3         33        105         23 |       164
  Key Stage V |         2         34         67         18 |       121
--------------+--------------------------------------------+----------
        Total |         9        143        382        133 |       667
```

Looking at teachers' rating of SIM instructions by key stage, it shows that consistently in all key stages, the mode is 3, which is "effective."

Table 172: SIM teachers' rating of SIM instructions frequency distribution, by school type

```
. tabulate school q22

           |                         q22
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          1          3          1 |         5
       HSS |         3         61        171         55 |       290
       LSS |         0         10         28         14 |        52
       MSS |         6         55        104         30 |       195
        PS |         0         16         76         33 |       125
-----------+--------------------------------------------+----------
     Total |         9        143        382        133 |       667
```

Likewise, looking at teachers' rating of SIM instructions by school type, it shows that consistently in all school types, the mode is 3, which is "effective."

Table 173: Result of the SIM teachers' rating of SIM instructions median calculation

```
. tabstat q22, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q22 |       667         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM teacher respondents found SIM instructions "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 174: Result of the SIM teachers' measure of consensus on SIM instructions rating

```
. cns q22 , min(1) max(4)

Consensus Measure for q22
Cns(X) = .72603986
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM instructions rating of SIM teachers, it is 0.7260.

Inferential Analysis – Statistical Significance Testing

Table 175: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q22 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      515    179800.5      111389
    Negative |      152     42977.5      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -3074286.4
Adjustment for zeros          0
                     ----------
Adjusted variance      21709766

H0: q22 = 2.5
         z = 14.683
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 77.2% of SIM teachers surveyed think that SIM instructions was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = “ineffective” and 3 = “effective.”

Ho: Our null hypothesis is that the SIM teacher population’s true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population’s true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.683, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 14.683 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 14.683/SQRT(667) = 0.57. This, according to Bartz (1999) is moderate effect size.

Gender difference in SIM teachers' rating of SIM instructions

Table 176: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q22, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       93917       89178
        Male |      400      128861      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.22e+06
                     ----------
Adjusted variance    4722665.14

H0: q22(gender==Female) = q22(gender==Male)
         z =  2.181
Prob > |z| = 0.0292
Exact prob = 0.0288
```

There is evidence for statistically significant difference between SIM instructions rating between female teachers and male teachers (p-value = 0.0288 < alpha = 0.05), which means female teachers and male teachers rated SIM instructions differently effective.

Evidence on SIM Teachers' Perception on SIM Instructions

There is statistically significant evidence (p = 0.0000) that the majority 77.2% of SIM teachers found instructions in SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 14.683, p = 0.0000, with a moderate effect size (*r* = 0.57).

Analyzing Teachers' Perception on Graphics in SIM Booklets

The Ministry of Education was interested to know how SIM teachers found graphics in the SIM booklets. To investigate this, ***Figure 37*** shows the results of SIM teachers' perception on graphics in the SIM booklets.

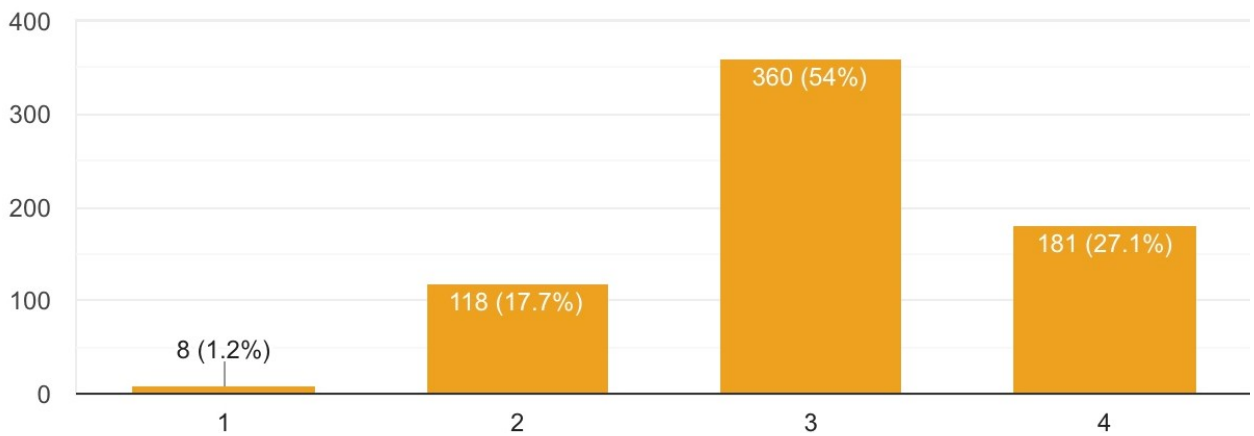


Figure 37*: Results of "Rate how did you find graphics in the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 37*** the 81.1% of the SIM teacher respondents rated the SIM graphics "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 177: Results of the SIM teachers' rating of SIM graphics frequency distribution

```
. tabulate q24

        q24 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          8        1.20        1.20
          2 |        118       17.69       18.89
          3 |        360       53.97       72.86
          4 |        181       27.14      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM teacher respondents of 81.1% chose "effective" or "extremely effective."

Table 178: SIM teachers' rating of SIM graphics frequency distribution, by age group

```
. tabulate age_group q24
```

Age_Group	q24: 1	2	3	4	Total
(20-24)	0	0	2	3	5
(25-29)	4	22	114	48	188
(30-34)	3	49	99	52	203
(35-39)	0	30	74	50	154
(40-44)	1	8	41	15	65
(45-49)	0	7	19	10	36
(50-54)	0	2	7	3	12
(55-59)	0	0	4	0	4
Total	8	118	360	181	667

Looking at teachers' rating of SIM graphics by age group, it shows that in majority age groups the mode is 3, which is "effective." Interestingly, the youngest age group of 20-24 year old rated SIM graphics "extremely effective" as they have mode of 4.

Table 179: SIM teachers' rating of SIM graphics frequency distribution, by key stage

```
. tabulate key_stage q24
```

Key_Stage	q24: 1	2	3	4	Total
Key Stage I	1	25	67	45	138
Key Stage II	1	16	98	50	165
Key Stage III	1	16	42	20	79
Key Stage IV	4	27	89	44	164
Key Stage V	1	34	64	22	121
Total	8	118	360	181	667

Similarly, looking at teachers' rating of SIM graphics by key stage, it shows that all key stages have the mode as 3, which is "effective."

Table 180: SIM teachers' rating of SIM graphics frequency distribution, by school type

```
. tabulate school q24

           |                    q24
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         0          1          2          2 |         5
       HSS |         2         55        166         67 |       290
       LSS |         0          6         25         21 |        52
       MSS |         5         47         90         53 |       195
        PS |         1          9         77         38 |       125
-----------+--------------------------------------------+----------
     Total |         8        118        360        181 |       667
```

Likewise, looking at teachers' rating of SIM graphics by school type, it shows that almost all school types rated SIM graphics as "effective" with mode of 3.

Table 181: Result of the SIM teachers' rating of SIM graphics median calculation

```
. tabstat q24, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q24 |       667         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM teacher respondents found SIM graphics "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 182: Result of the SIM teachers' measure of consensus on SIM graphics rating

```
. cns q24 , min(1) max(4)

Consensus Measure for q24
Cns(X) = .70339456
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM graphics rating of SIM teachers, it is 0.7034.

Inferential Analysis – Statistical Significance Testing

Table 183: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q24 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      541      189933      111389
    Negative |      126       32845      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -2415957.4
Adjustment for zeros          0
                     ----------
Adjusted variance      22368095

H0: q24 = 2.5
         z = 16.607
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 81.1% of SIM teachers surveyed think that SIM graphics was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 16.607, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 16.607 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 16.607/SQRT(667) = 0.64. This, according to Bartz (1999), is strong effect size.

Gender difference in SIM teachers' rating of SIM graphics

Table 184: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q24, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       92865       89178
        Male |      400      129913      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.09e+06
                     ----------
Adjusted variance    4858728.15

H0: q24(gender==Female) = q24(gender==Male)
         z =  1.673
Prob > |z| = 0.0944
Exact prob = 0.0954
```

There is no evidence for statistically significant difference between SIM graphics rating between female teachers and male teachers (p-value = 0.0954 > alpha = 0.05), which means both female teachers and male teachers found SIM graphics equally effective.

Evidence on SIM Teachers' Perception of SIM Graphics

There is statistically significant evidence (p = 0.0000) that the majority 81.1% of SIM teachers, both female teachers and male teachers, found graphics in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 16.607, p = 0.0000, with a strong effect size (*r* = 0.64).

Analyzing Teachers' Perception on Activities in SIM Booklets

The Ministry of Education was interested to know how SIM teachers found activities in the SIM booklets. To investigate this, ***Figure 38*** shows the results of SIM teachers' perception on activities in the SIM booklets.

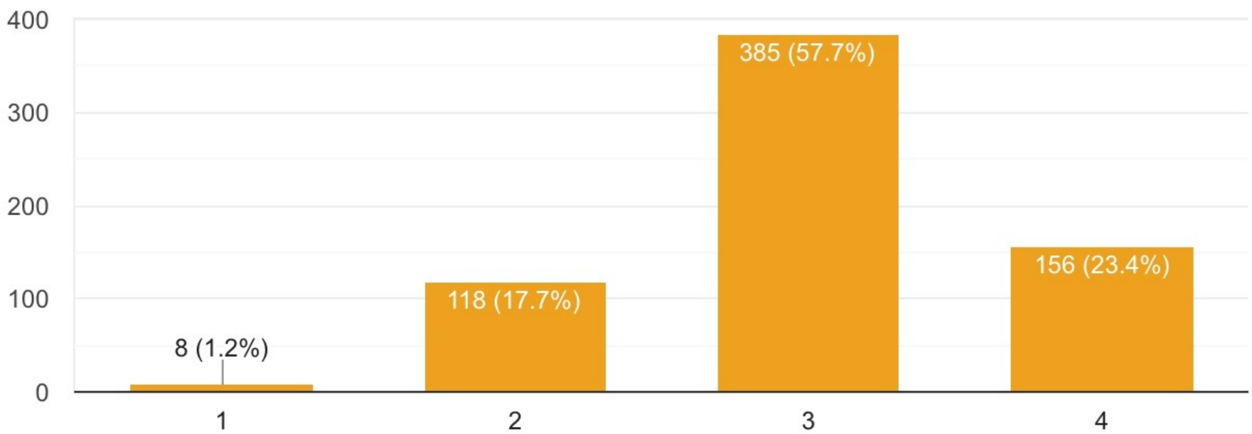


Figure 38*: Results of "Rate how did you find activities in the SIM materials" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 38*** the 81.1% of the SIM teacher respondents rated the SIM activities "effective" or "extremely effective."

Descriptive Analysis – Measure of Central Tendency

Table 185: Results of the SIM teachers' rating of SIM activities frequency distribution

```
. tabulate q25

        q25 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          8        1.20        1.20
          2 |        118       17.69       18.89
          3 |        385       57.72       76.61
          4 |        156       23.39      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM teacher respondents of 81.1% chose "effective" or "extremely effective."

Table 186: SIM teachers' rating of SIM activities frequency distribution, by age group

```
. tabulate age_group q25
```

Age_Group	q25 1	2	3	4	Total
(20-24)	0	0	4	1	5
(25-29)	2	40	109	37	188
(30-34)	5	34	121	43	203
(35-39)	0	26	80	48	154
(40-44)	1	8	39	17	65
(45-49)	0	8	20	8	36
(50-54)	0	1	10	1	12
(55-59)	0	1	2	1	4
Total	8	118	385	156	667

Looking at teachers' rating of SIM activities by age group, it shows consistently that all age groups have mode 3, which is "effective."

Table 187: SIM teachers' rating of SIM activities frequency distribution, by key stage

```
. tabulate key_stage q25
```

Key_Stage	q25 1	2	3	4	Total
Key Stage I	1	25	73	39	138
Key Stage II	0	16	108	41	165
Key Stage III	1	12	42	24	79
Key Stage IV	2	32	95	35	164
Key Stage V	4	33	67	17	121
Total	8	118	385	156	667

Similarly, looking at teachers' rating of SIM activities by key stage, it shows consistently that all key stages have mode 3, which is "effective."

Table 188: SIM teachers' rating of SIM activities frequency distribution, by school type

```
. tabulate school q25
```

School	q25: 1	2	3	4	Total
ECR	0	1	2	2	5
HSS	5	59	167	59	290
LSS	0	5	28	19	52
MSS	3	42	108	42	195
PS	0	11	80	34	125
Total	8	118	385	156	667

Likewise, looking at teachers' rating of SIM activities by school type, it shows all school types rated SIM activities as "effective" with mode of 3 while interestingly ECR bimodal ratings of "effective" and "extremely effective."

Table 189: Result of the SIM teachers' rating of SIM activities median calculation

```
. tabstat q25, stat(count p50 min max)
```

Variable	N	p50	Min	Max
q25	667	3	1	4

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM teacher respondents found SIM activities "effective" or "extremely effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 190: Result of the SIM teachers' measure of consensus on SIM activities rating

```
. cns q25 , min(1) max(4)

Consensus Measure for q25
Cns(X) = .73217896
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM activities rating of SIM teachers, it is 0.7322.

Inferential Analysis – Statistical Significance Testing

Table 191: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q25 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      541      188358      111389
    Negative |      126       34420      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -2743204.3
Adjustment for zeros          0
                     ----------
Adjusted variance      22040848

H0: q25 = 2.5
         z = 16.395
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 81.1% of SIM teachers surveyed think that SIM activities was effective or extremely effective. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 16.395, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 16.395 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 16.395/SQRT(667) = 0.63. This, according to Bartz (1999), is strong effect size.

Gender difference in SIM teachers' rating of SIM activities

Table 192: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q25, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       92533       89178
        Male |      400      130245      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.25e+06
                     ----------
Adjusted variance    4692892.21

H0: q25(gender==Female) = q25(gender==Male)
         z =  1.549
Prob > |z| = 0.1214
Exact prob = 0.1222
```

There is no evidence for statistically significant difference between SIM activities rating between female teachers and male teachers (p-value = 0.1222 > alpha = 0.05), which means both female teachers and male teachers found SIM activities equally effective.

Evidence on SIM Teachers' Perception of SIM Activities

There is statistically significant evidence (p = 0.0000) that the majority 81.1% of SIM teachers, both female teachers and male teachers, found activities in the SIM booklets effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 16.395, p = 0.0000, with a strong effect size (r = 0.63).

Effectiveness of SIM Learning

Analyzing Teachers' Perception on SIM Learning in Increasing Knowledge

The Ministry of Education was interested to know how SIM teachers found SIM learning in increasing knowledge. To investigate this, ***Figure 39*** shows the results of SIM teachers' perception on increasing knowledge during SIM learning.

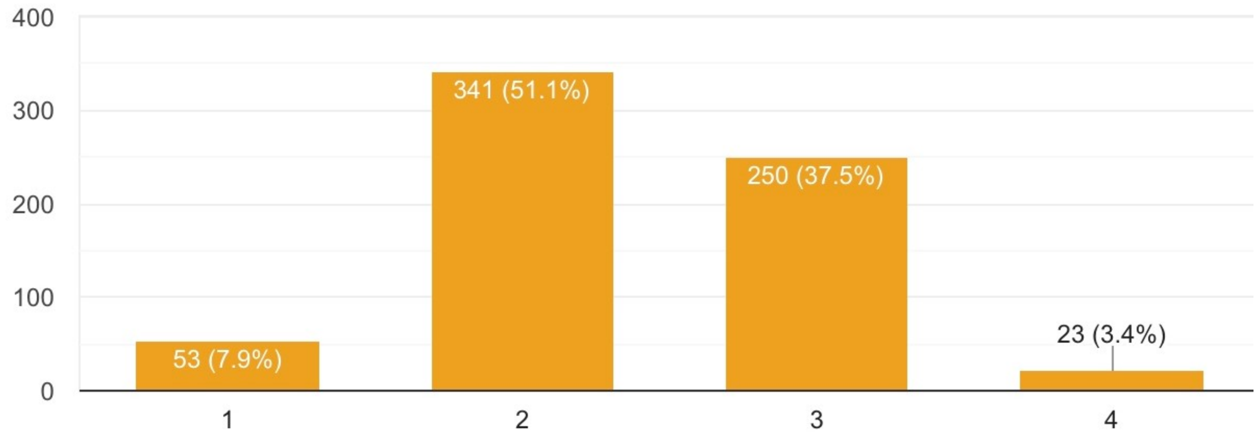


Figure 39*: Results of "Rate the effectiveness of SIM-learning in terms of increasing knowledge" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 39*** only 40.9% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in increasing knowledge.

Descriptive Analysis – Measure of Central Tendency

Table 193: Results of the SIM teachers' rating of SIM learning in increasing knowledge

```
. tabulate q7

         q7 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         53        7.95        7.95
          2 |        341       51.12       59.07
          3 |        250       37.48       96.55
          4 |         23        3.45      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "ineffective." The total SIM teacher respondents of only 40.9% chose "effective" or "extremely effective" for SIM learning in increasing knowledge.

Table 194: SIM teachers' rating of SIM learning in increasing knowledge, by age group

```
. tabulate age_group q7

           |                         q7
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         1          3          1          0 |         5
   (25-29) |        11        112         60          5 |       188
   (30-34) |        22         97         74         10 |       203
   (35-39) |        14         76         60          4 |       154
   (40-44) |         3         29         32          1 |        65
   (45-49) |         1         19         14          2 |        36
   (50-54) |         1          4          6          1 |        12
   (55-59) |         0          1          3          0 |         4
-----------+--------------------------------------------+----------
     Total |        53        341        250         23 |       667
```

Looking at teachers' rating of SIM learning in increasing knowledge, by age group, it shows that in majority of age groups, the mode is 2, which is "ineffective."

Table 195: SIM teachers' rating of SIM learning in increasing knowledge, by key stage

```
. tabulate key_stage q7

              |                         q7
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         7         70         55          6 |       138
 Key Stage II |        12         72         74          7 |       165
Key Stage III |        11         43         23          2 |        79
 Key Stage IV |        14         87         57          6 |       164
  Key Stage V |         9         69         41          2 |       121
--------------+--------------------------------------------+----------
        Total |        53        341        250         23 |       667
```

Similarly, looking at teachers' rating of SIM learning in increasing knowledge, by key stage, it shows that in majority of key stages, the mode is 2, which is "ineffective."

Table 196: SIM teachers' rating of SIM learning in increasing knowledge, by school type

```
. tabulate school q7

           |                  q7
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1          2          2          0 |         5
       HSS |        19        158        103         10 |       290
       LSS |         2         28         21          1 |        52
       MSS |        25        106         61          3 |       195
        PS |         6         47         63          9 |       125
-----------+--------------------------------------------+----------
     Total |        53        341        250         23 |       667
```

Likewise, looking at teachers' rating of SIM learning in increasing knowledge, by school type, it shows that in majority of school types especially higher level schools, the mode is 2, which is "ineffective."

Table 197: Median of the SIM teachers' rating of SIM learning in increasing knowledge

```
. tabstat q7, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q7 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in increasing knowledge.

Descriptive Analysis - Measure of Dispersion

Table 198: SIM teachers' measure of consensus on SIM learning in increasing knowledge

```
. cns q7 , min(1) max(4)

Consensus Measure for q7
Cns(X) = .66701141
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in increasing knowledge, it is 0.6670.

Inferential Analysis – Statistical Significance Testing

Table 199: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q7 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      273     88478.5      111389
    Negative |      394    134299.5      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -4309653.8
Adjustment for zeros          0
                     ----------
Adjusted variance      20474399

H0: q7 = 2.5
         z = -5.063
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that only 40.9% of SIM teachers surveyed think that SIM learning was effective or extremely effective in increasing knowledge. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -5.063, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -5.063 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -5.063/SQRT(667) = -0.20. Neglecting negative sign, this, according to Bartz (1999), is very low effect size.

Gender difference in SIM teachers' perception of SIM learning in increasing knowledge

Table 200: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q7, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267     85547.5       89178
        Male |      400    137230.5      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.11e+06
                     ----------
Adjusted variance    4834514.21

H0: q7(gender==Female) = q7(gender==Male)
         z = -1.651
Prob > |z| = 0.0987
Exact prob = 0.0993
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.0993 > alpha = 0.05) on perception of SIM learning in increasing knowledge, which means female teachers and male teachers rated similar on SIM learning effectiveness in increasing knowledge.

Evidence on SIM Teachers' Perception of SIM Learning in Increasing Knowledge

There is statistically significant evidence (p = 0.0000) that only minority 40.9% of SIM teachers, both female teachers and male teachers, found SIM learning effective in increasing knowledge. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -5.063, p = 0.0000, with a very low effect size (r = 0.20).

Analyzing Teachers' Perception on SIM Learning in Increasing Skills

The Ministry of Education was interested to know how SIM teachers found SIM learning in increasing skills. To investigate this, ***Figure 40*** shows the results of SIM teachers' perception on increasing skills during SIM learning.

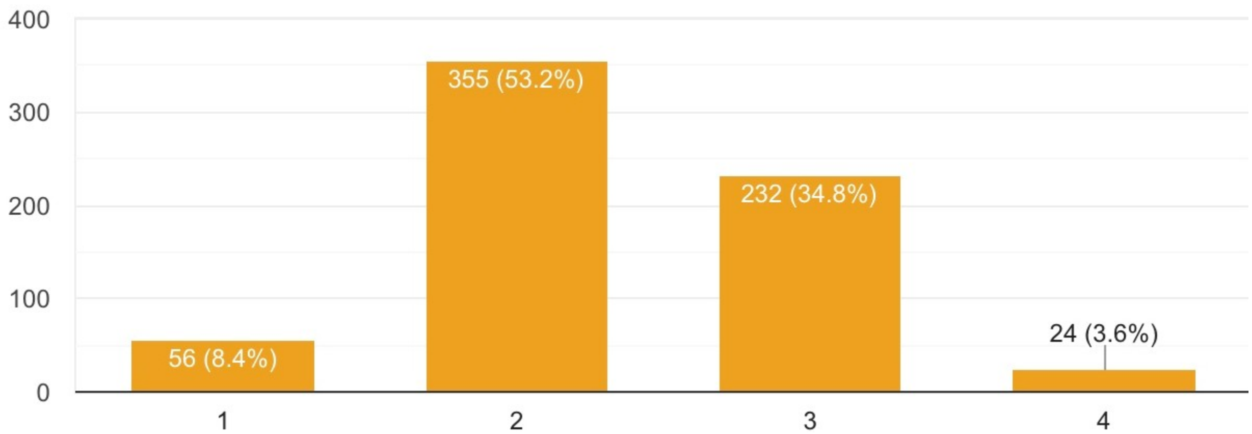


Figure 40*: Results of "Rate the effectiveness of SIM-learning in terms of increasing skills" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 40*** the 38.4% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in increasing skills.

Descriptive Analysis – Measure of Central Tendency

Table 201: Results of the SIM teachers' rating of SIM learning in increasing skills

```
. tabulate q8

         q8 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         56        8.40        8.40
          2 |        355       53.22       61.62
          3 |        232       34.78       96.40
          4 |         24        3.60      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "ineffective." The total SIM teacher respondents of 38.4% chose "effective" or "extremely effective" for SIM learning in increasing skills.

Table 202: SIM teachers' rating of SIM learning in increasing skills, by age group

```
. tabulate age_group q8
```

Age_Group	q8 1	2	3	4	Total
(20-24)	1	3	1	0	5
(25-29)	14	109	59	6	188
(30-34)	21	100	75	7	203
(35-39)	13	77	58	6	154
(40-44)	4	40	19	2	65
(45-49)	2	16	17	1	36
(50-54)	1	7	2	2	12
(55-59)	0	3	1	0	4
Total	56	355	232	24	667

Looking at teachers' rating of SIM learning in increasing skills, by age group, it shows that in almost all age groups, the mode is 2, which is "ineffective."

Table 203: SIM teachers' rating of SIM learning in increasing skills, by key stage

```
. tabulate key_stage q8
```

Key_Stage	q8 1	2	3	4	Total
Key Stage I	9	74	46	9	138
Key Stage II	12	78	70	5	165
Key Stage III	7	44	25	3	79
Key Stage IV	15	93	49	7	164
Key Stage V	13	66	42	0	121
Total	56	355	232	24	667

Looking at teachers' rating of SIM learning in increasing skills, by key stage, it shows that all key stages have mode as 2, which is "ineffective."

Table 204: SIM teachers' rating of SIM learning in increasing skills, by school type

```
. tabulate school q8

           |                     q8
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1          3          1          0 |         5
       HSS |        22        166         94          8 |       290
       LSS |         4         30         16          2 |        52
       MSS |        23        104         61          7 |       195
        PS |         6         52         60          7 |       125
-----------+--------------------------------------------+----------
     Total |        56        355        232         24 |       667
```

Looking at teachers' rating of SIM learning in increasing skills, by school type, it shows that all school types except PS have the mode as 2, which is "ineffective." PS rated SIM learning "effective" in increasing skills as it has its mode as 3.

Table 205: Median of the SIM teachers' rating of SIM learning in increasing skills

```
. tabstat q8, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q8 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in increasing skills.

Descriptive Analysis - Measure of Dispersion

Table 206: SIM teachers' measure of consensus on SIM learning in increasing skills

```
. cns q8 , min(1) max(4)

Consensus Measure for q8
Cns(X) = .66988918
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in increasing skills, it is 0.6699.

Inferential Analysis – Statistical Significance Testing

Table 207: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      256       83268      111389
    Negative |      411      139510      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -4224444.5
Adjustment for zeros          0
                     ----------
Adjusted variance      20559608

H0: q8 = 2.5
         z = -6.202
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that only 38.4% of SIM teachers surveyed think that SIM learning was effective or extremely effective in increasing skills. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -6.202, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -6.202 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -6.202/SQRT(667) = -0.24. Neglecting negative sign, this, according to Bartz (1999) is low effect size.

Gender difference in SIM teachers' perception of SIM learning in increasing skills

Table 208: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q8, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       89417       89178
        Male |      400      133361      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.15e+06
                     ----------
Adjusted variance    4794892.58

H0: q8(gender==Female) = q8(gender==Male)
         z =  0.109
Prob > |z| = 0.9131
Exact prob = 0.9145
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.9145 > alpha = 0.05) on perception of SIM learning in increasing skills, which means female teachers and male teachers rated similar on SIM learning effectiveness in increasing skills.

Evidence on SIM Teachers' Perception of SIM Learning in Increasing Skills

There is statistically significant evidence (p = 0.0000) that only minority 38.4% of SIM teachers, both female teachers and male teachers, found SIM learning effective in increasing skills. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -6.202, p = 0.0000, with a low effect size (*r* = 0.24).

Analyzing Teachers' Perception on SIM Learning in Imparting Values

The Ministry of Education was interested to know how SIM teachers found SIM learning in imparting values. To investigate this, ***Figure 41*** shows the results of SIM teachers' perception on imparting values during SIM learning.

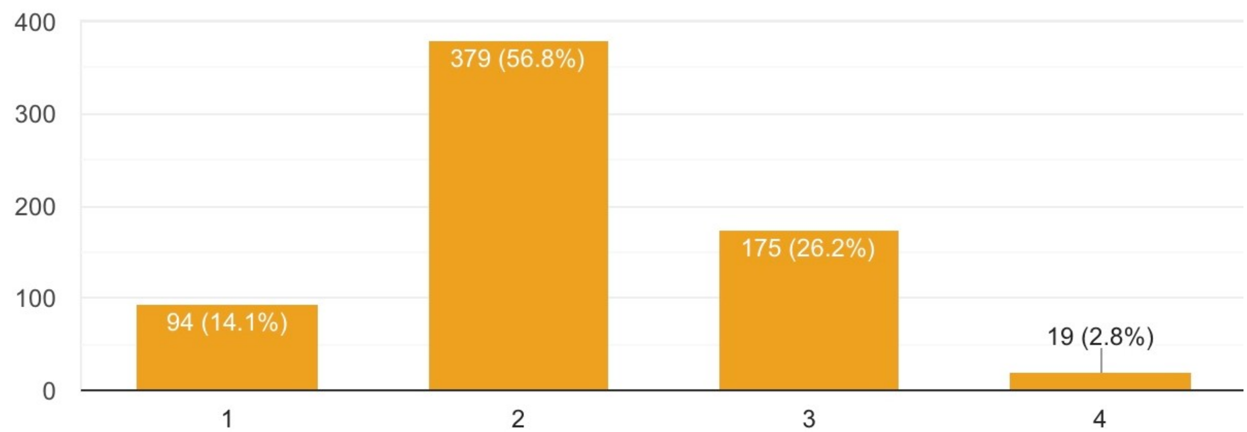


Figure 41*: Results of "Rate the effectiveness of SIM-learning in terms of imparting values" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 41*** the 29.0% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in imparting values.

Descriptive Analysis – Measure of Central Tendency

Table 209: Results of the SIM teachers' rating of SIM learning in imparting values

```
. tabulate q9

         q9 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         94       14.09       14.09
          2 |        379       56.82       70.91
          3 |        175       26.24       97.15
          4 |         19        2.85      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 3, which is "effective." The total SIM teacher respondents of only 29.0% chose "effective" or "extremely effective" for SIM learning in imparting values.

Table 210: SIM teachers' rating of SIM learning in imparting values, by age group

```
. tabulate age_group q9
```

Age_Group	q9 1	2	3	4	Total
(20-24)	1	4	0	0	5
(25-29)	29	112	45	2	188
(30-34)	28	109	59	7	203
(35-39)	21	87	40	6	154
(40-44)	8	39	16	2	65
(45-49)	5	19	10	2	36
(50-54)	2	6	4	0	12
(55-59)	0	3	1	0	4
Total	94	379	175	19	667

Looking at teachers' rating of SIM learning in imparting values, by age group, it shows that in all age groups, the mode is 2, which is "ineffective."

Table 211: SIM teachers' rating of SIM learning in imparting values, by key stage

```
. tabulate key_stage q9
```

Key_Stage	q9 1	2	3	4	Total
Key Stage I	9	75	51	3	138
Key Stage II	18	96	44	7	165
Key Stage III	18	42	17	2	79
Key Stage IV	28	97	33	6	164
Key Stage V	21	69	30	1	121
Total	94	379	175	19	667

Similarly, looking at teachers' rating of SIM learning in imparting values, by key stage, it shows that in all key stages the mode is 2, which is "ineffective."

Table 212: SIM teachers' rating of SIM learning in imparting values, by school type

School	q9: 1	2	3	4	Total
ECR	1	2	2	0	5
HSS	44	169	71	6	290
LSS	8	28	13	3	52
MSS	33	114	42	6	195
PS	8	66	47	4	125
Total	94	379	175	19	667

Teachers' rating of SIM learning in imparting values by school type shows that in all school types the mode is 2, which is "ineffective."

Table 213: Median of the SIM teachers' rating of SIM learning in imparting values

```
. tabstat q9, stat(count p50 min max)
```

Variable	N	p50	Min	Max
q9	667	2	1	4

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in imparting values.

Descriptive Analysis - Measure of Dispersion

Table 214: SIM teachers' measure of consensus on SIM learning in imparting values

```
. cns q9 , min(1) max(4)

Consensus Measure for q9
Cns(X) = .68876827
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in imparting values, it is 0.6888.

Inferential Analysis – Statistical Significance Testing

Table 215: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q9 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      194     60171.5      111389
    Negative |      473    162606.5      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -3572368.6
Adjustment for zeros          0
                     ----------
Adjusted variance      21211684

H0: q9 = 2.5
         z = -11.121
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 29.0% of SIM teachers surveyed think that SIM learning was effective or extremely effective in imparting values. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -11.121, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -11.121 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -11.121/SQRT(667) = -0.43. This, according to Bartz (1999) ,is moderate effect size.

Gender difference in SIM teachers' perception of SIM learning in imparting values

Table 216: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q9, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann,ÄìWhitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       86047       89178
        Male |      400      136731      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.21e+06
                     ----------
Adjusted variance    4730354.27

H0: q9(gender==Female) = q9(gender==Male)
         z = -1.440
Prob > |z| = 0.1500
Exact prob = 0.1506
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.1506 > alpha = 0.05) on perception of SIM learning in imparting values, which means female teachers and male teachers rated similarly on SIM learning effectiveness in imparting values.

Evidence on SIM Teachers' Perception of SIM Learning in Imparting Values

There is statistically significant evidence (p = 0.0000) that only minority 29.0% of SIM teachers, both female teachers and males teachers, found SIM learning effective in imparting values. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -11.121, p = 0.0000, with a moderate effect size (*r* = 0.43).

Analyzing Teachers' Perception on SIM Learning in Improving Attitudes

The Ministry of Education was interested to know how SIM teachers found SIM learning in improving attitudes. To investigate this, ***Figure 42*** shows the results of SIM teachers' perception on improving attitudes during SIM learning in comparison to classroom learning.

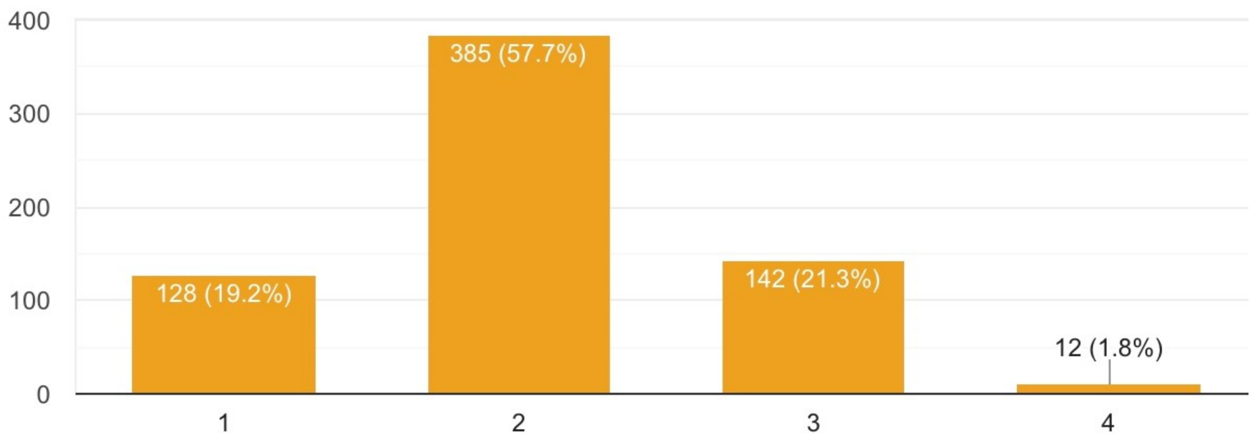


Figure 42*: Results of "Rate the effectiveness of SIM-learning in terms of improving attitudes" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 42*** only 23.1% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in improving attitudes.

Descriptive Analysis – Measure of Central Tendency

Table 217: Results of the SIM teachers' rating of SIM learning in improving attitudes

```
. tabulate q10

        q10 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        128       19.19       19.19
          2 |        385       57.72       76.91
          3 |        142       21.29       98.20
          4 |         12        1.80      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "ineffective." The total SIM teacher respondents of only 23.1% chose "effective" or "extremely effective" for SIM learning in improving attitudes.

Table 218: SIM teachers' rating of SIM learning in improving attitudes, by age group

```
. tabulate age_group q10

           |                    q10
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         2          1          2          0 |         5
   (25-29) |        41        116         30          1 |       188
   (30-34) |        35        120         43          5 |       203
   (35-39) |        34         79         39          2 |       154
   (40-44) |         8         41         14          2 |        65
   (45-49) |         4         22         10          0 |        36
   (50-54) |         4          4          2          2 |        12
   (55-59) |         0          2          2          0 |         4
-----------+--------------------------------------------+----------
     Total |       128        385        142         12 |       667
```

Looking at teachers' rating of SIM learning in improving attitudes, by age group, it shows that majority rated SIM learning ineffective in improving attitudes.

Table 219: SIM teachers' rating of SIM learning in improving attitudes, by key stage

```
. tabulate key_stage q10

              |                    q10
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |        16         87         31          4 |       138
 Key Stage II |        26         96         40          3 |       165
Key Stage III |        22         39         15          3 |        79
 Key Stage IV |        42         91         29          2 |       164
  Key Stage V |        22         72         27          0 |       121
--------------+--------------------------------------------+----------
        Total |       128        385        142         12 |       667
```

Looking at teachers' rating of SIM learning in improving attitudes, by key stage, it shows that all key stages have the mode as 2, which is "ineffective".

Table 220: SIM teachers' rating of SIM learning in improving attitudes, by school type

```
. tabulate school q10

           |                     q10
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1          3          1          0 |         5
       HSS |        58        171         58          3 |       290
       LSS |        10         30         10          2 |        52
       MSS |        50        103         39          3 |       195
        PS |         9         78         34          4 |       125
-----------+--------------------------------------------+----------
     Total |       128        385        142         12 |       667
```

Like by key stage, teachers' rating of SIM learning in improving attitudes by school type is also 2 for all school types, which is "ineffective."

Table 221: Median of the SIM teachers' rating of SIM learning in improving attitudes

```
. tabstat q10, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q10 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in improving attitudes.

Descriptive Analysis - Measure of Dispersion

Table 222: SIM teachers' measure of consensus on SIM learning in improving attitudes

```
. cns q10 , min(1) max(4)

Consensus Measure for q10
Cns(X) = .72079116
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in improving attitudes, it is 0.7208.

Inferential Analysis – Statistical Significance Testing

Table 223: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q10 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      154       44658      111389
    Negative |      513      178120      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -3106385.8
Adjustment for zeros          0
                     ----------
Adjusted variance      21677667

H0: q10 = 2.5
         z = -14.332
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that only 23.1% of SIM teachers surveyed think that SIM learning was effective or extremely effective in improving attitudes. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = “ineffective” and 3 = “effective.”

Ho: Our null hypothesis is that the SIM teacher population’s true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population’s true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -14.332, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -14.332 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -14.332/SQRT(667) = -0.56. Neglecting negative sign, this, according to Bartz (1999), is moderate effect size.

Gender difference in SIM teachers' perception of SIM learning in improving attitudes

Table 224: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       86857       89178
        Male |      400      135921      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.24e+06
                     ----------
Adjusted variance    4702464.79

H0: q10(gender==Female) = q10(gender==Male)
         z = -1.070
Prob > |z| = 0.2845
Exact prob = 0.2863
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.2863 > alpha = 0.05) on perception of SIM learning in improving attitudes, which means female teachers and male teachers rated similar on SIM learning effectiveness in improving attitudes.

Evidence on SIM Teachers' Perception of SIM Learning in Improving Attitudes

There is statistically significant evidence (p = 0.0000) that only minority 23.1% of SIM teachers, both female teachers and male teachers, found SIM learning effective in improving attitudes. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -14.332, p = 0.0000, with a moderate effect size (*r* = 0.56).

Analyzing Teachers' Perception on SIM Learning in Understanding English

The Ministry of Education was interested to know how SIM teachers found SIM learning in understanding English. To investigate this, ***Figure 43*** shows the results of SIM teachers' perception on understanding English during SIM learning.

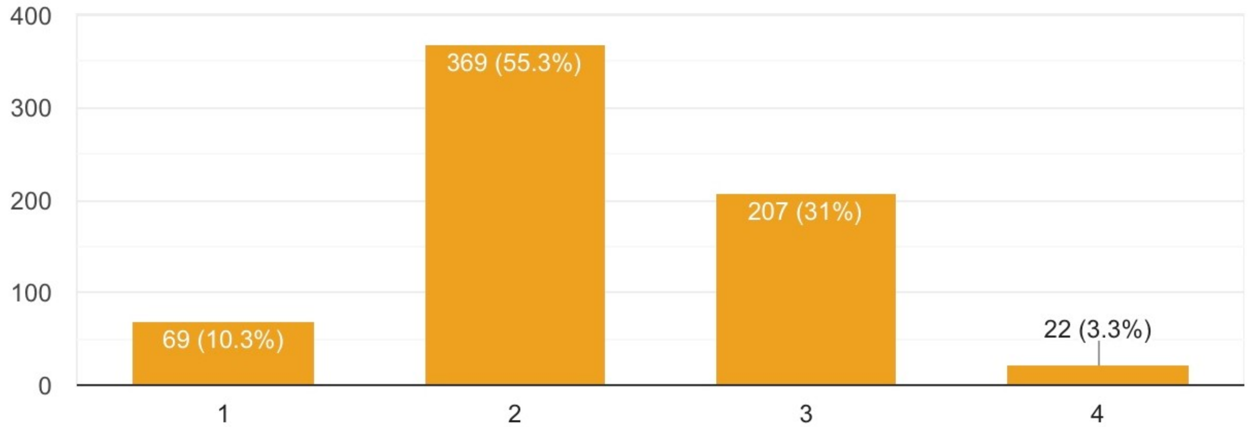


Figure 43*: Results of "Rate the effectiveness of SIM-learning in terms of understanding English subject" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 43*** only 34.3% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in understanding English.

Descriptive Analysis – Measure of Central Tendency

Table 225: Results of the SIM teachers' rating of SIM learning in understanding English

```
. tabulate q11

        q11 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         69       10.34       10.34
          2 |        369       55.32       65.67
          3 |        207       31.03       96.70
          4 |         22        3.30      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "ineffective." The total SIM teacher respondents of only 34.3% chose "effective" or "extremely effective" for SIM learning in understanding English.

Table 226: SIM teachers' rating of SIM learning in understanding English, by age group

```
. tabulate age_group q11

           |                         q11
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         1          1          3          0 |         5
   (25-29) |        14        118         53          3 |       188
   (30-34) |        25        108         65          5 |       203
   (35-39) |        17         83         46          8 |       154
   (40-44) |         4         35         25          1 |        65
   (45-49) |         5         19         10          2 |        36
   (50-54) |         2          5          3          2 |        12
   (55-59) |         1          0          2          1 |         4
-----------+--------------------------------------------+----------
     Total |        69        369        207         22 |       667
```

Looking at teachers' rating of SIM learning in understanding English, by age group, it shows that the majority of the age groups except 20-24 and 55-59 have the mode as 2, which is "ineffective". But the youngest age group and the oldest age group have mode as 3, which is "effective."

Table 227: SIM teachers' rating of SIM learning in understanding English, by key stage

```
. tabulate key_stage q11

              |                         q11
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         7         88         39          4 |       138
 Key Stage II |        16         87         56          6 |       165
Key Stage III |        16         39         22          2 |        79
 Key Stage IV |        17         91         52          4 |       164
  Key Stage V |        13         64         38          6 |       121
--------------+--------------------------------------------+----------
        Total |        69        369        207         22 |       667
```

Looking at teachers' rating of SIM learning in understanding English by key stage, it shows that all key stages have the mode as 2, which is "ineffective".

Table 228: SIM teachers' rating of SIM learning in understanding English, by school type

```
. tabulate school q11

           |                          q11
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1          3          1          0 |         5
       HSS |        33        156         91         10 |       290
       LSS |         7         30         13          2 |        52
       MSS |        24        110         56          5 |       195
        PS |         4         70         46          5 |       125
-----------+--------------------------------------------+----------
     Total |        69        369        207         22 |       667
```

Looking at teachers' rating of SIM learning in understanding English by school type, it shows school types have mode as 2, which is "ineffective."

Table 229: Median of the SIM teachers' rating of SIM learning in understanding English

```
. tabstat q11, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q11 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in understanding English.

Descriptive Analysis - Measure of Dispersion

Table 230: SIM teachers' measure of consensus on SIM learning in understanding English

```
. cns q11 , min(1) max(4)

Consensus Measure for q11
Cns(X) = .67641023
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in understanding English, it is 0.6764.

Inferential Analysis – Statistical Significance Testing

Table 231: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q11 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      229     73403.5      111389
    Negative |      438    149374.5      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -3996997.5
Adjustment for zeros          0
                     ----------
Adjusted variance      20787055

H0: q11 = 2.5
         z = -8.331
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 34.3% of SIM teachers surveyed think that SIM learning was effective or extremely effective in understanding English. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -8.331, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -8.331 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -8.331/SQRT(667) = -0.32. Neglecting negative sign, this, according to Bartz (1999), is low effect size.

Gender difference in SIM teachers' perception of SIM learning in understanding English

Table 232: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q11, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       87569       89178
        Male |      400      135209      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.19e+06
                     ----------
Adjusted variance    4754084.72

H0: q11(gender==Female) = q11(gender==Male)
         z = -0.738
Prob > |z| = 0.4605
Exact prob = 0.4615
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.4615 >alpha = 0.05) on perception of SIM learning in understanding English, which means female teachers and male teachers rated similarly on SIM learning effectiveness in understanding English.

Evidence on SIM Teachers' Perception of SIM Learning in Understanding English

There is statistically significant evidence (p = 0.0000) that only minority 34.3% of SIM teachers, both female teachers and male teachers, found SIM learning effective in understanding English. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -8.331, p = 0.0000, with low effect size (*r* = 0.32).

Analyzing Teachers' Perception on SIM Learning in Understanding Mathematics

The Ministry of Education was interested to know how SIM teachers found SIM learning in understanding Mathematics. To investigate this, ***Figure 44*** shows the results of SIM teachers' perception on understanding Mathematics during SIM learning in comparison to classroom learning.

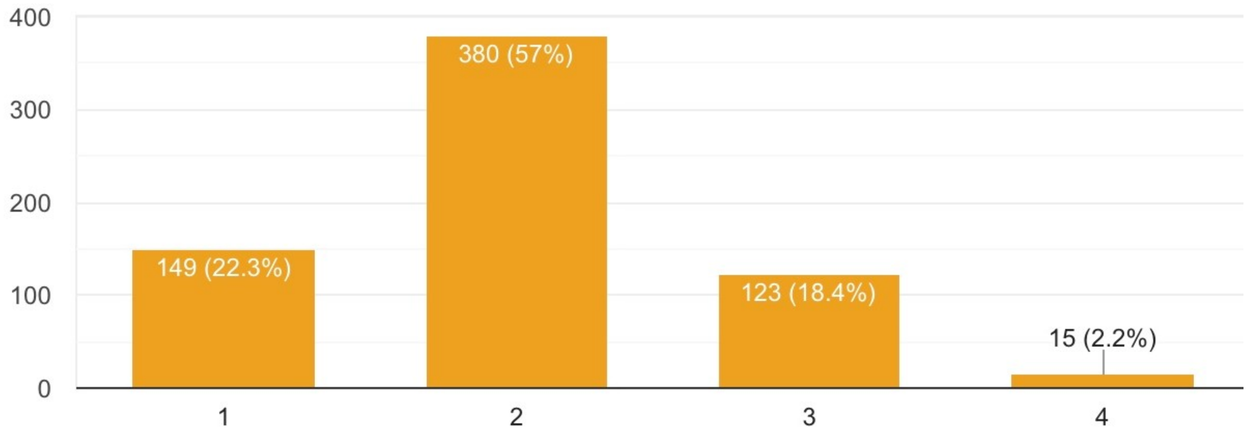


Figure 44*: Results of "Rate the effectiveness of SIM-learning in terms of understanding Mathematics subject" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 44*** only 20.6% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in understanding Mathematics.

Descriptive Analysis – Measure of Central Tendency

Table 233: Results of the SIM teachers' rating of SIM learning in understanding Mathematics

```
. tabulate q12

        q12 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        149       22.34       22.34
          2 |        380       56.97       79.31
          3 |        123       18.44       97.75
          4 |         15        2.25      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "ineffective." The total SIM teacher respondents of only 20.6% chose "effective" or "extremely effective" for SIM learning in understanding Mathematics.

Table 234: SIM teachers' rating of SIM learning in understanding Mathematics, by age group

```
. tabulate age_group q12
```

Age_Group	q12 1	2	3	4	Total
(20-24)	1	3	1	0	5
(25-29)	44	106	34	4	188
(30-34)	43	121	37	2	203
(35-39)	37	82	30	5	154
(40-44)	14	38	11	2	65
(45-49)	7	21	6	2	36
(50-54)	3	6	3	0	12
(55-59)	0	3	1	0	4
Total	149	380	123	15	667

Looking at teachers' rating of SIM learning in understanding Mathematics, by age group, it shows all age groups have the mode as 2, which is "ineffective".

Table 235: SIM teachers' rating of SIM learning in understanding Mathematics, by key stage

```
. tabulate key_stage q12
```

Key_Stage	q12 1	2	3	4	Total
Key Stage I	11	93	30	4	138
Key Stage II	23	95	43	4	165
Key Stage III	22	44	11	2	79
Key Stage IV	51	86	25	2	164
Key Stage V	42	62	14	3	121
Total	149	380	123	15	667

Looking at teachers' rating of SIM learning in understanding Mathematics by key stage, it shows that all key stages have the mode as 2, which is "ineffective".

Table 236: SIM teachers' rating of SIM learning in understanding Mathematics, by school type

```
. tabulate school q12

           |                    q12
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1          2          2          0 |         5
       HSS |        83        158         45          4 |       290
       LSS |         9         34          6          3 |        52
       MSS |        50        111         30          4 |       195
        PS |         6         75         40          4 |       125
-----------+--------------------------------------------+----------
     Total |       149        380        123         15 |       667
```

Looking at teachers' rating of SIM learning in understanding Mathematics by school type, it shows all school types have mode as 2, which is "ineffective."

Table 237: Median of the SIM teachers' rating of SIM learning in understanding Mathematics

```
. tabstat q12, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q12 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in understanding Mathematics.

Descriptive Analysis - Measure of Dispersion

Table 238: SIM teachers' measure of consensus on SIM learning in understanding Mathematics

```
. cns q12 , min(1) max(4)

Consensus Measure for q12
Cns(X) = .72418875
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in understanding Mathematics, it is 0.7242.

Inferential Analysis – Statistical Significance Testing

Table 239: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q12 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      138     39778.5      111389
    Negative |      529    182999.5      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -2743204.3
Adjustment for zeros          0
                     ----------
Adjusted variance      22040848

H0: q12 = 2.5
         z = -15.253
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that only 20.6% of SIM teachers surveyed think that SIM learning was effective or extremely effective in understanding Mathematics. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = - 15.253, p = 0.0000. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -15.253 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -15.253/SQRT(667) = -0.59 (we can ignore the negative sign). This, according to Bartz (1999), is moderate effect size.

Gender difference in SIM teachers' perception of SIM learning in understanding Mathematics

Table 240: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q12, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann,ÄìWhitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       86357       89178
        Male |      400      136421      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties   -1.20e+06
                     ----------
Adjusted variance    4742226.34

H0: q12(gender==Female) = q12(gender==Male)
         z = -1.295
Prob > |z| = 0.1952
Exact prob = 0.1954
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.1954 > alpha = 0.05) on perception of SIM learning in understanding Mathematics, which means female teachers and male teachers rated similar on SIM learning effectiveness in understanding Mathematics.

Evidence on SIM Teachers' Perception of SIM Learning in Understanding Mathematics

There is statistically significant evidence (p = 0.0000) that only minority 20.6% of SIM teachers found SIM learning effective in understanding Mathematics. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -15.253, p = 0.0000, with a moderate effect size (*r* = 0.59).

Analyzing Teachers' Perception on SIM Learning in Understanding Dzongkha

The Ministry of Education was interested to know how SIM teachers found SIM learning in understanding Dzongkha. To investigate this, ***Figure 45*** shows the results of SIM teachers' perception on understanding Dzongkha during SIM learning in comparison to classroom learning.

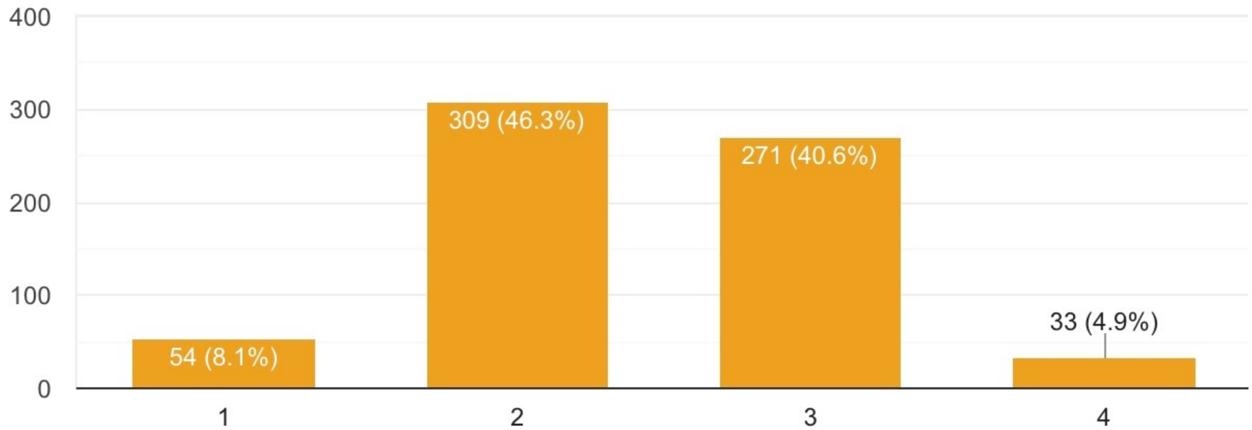


***Figure 45**: Results of "Rate the effectiveness of SIM-learning in terms of understanding Dzongkha subject" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 45*** only 45.5% of the SIM teacher respondents rated the SIM learning "effective" or "extremely effective" in understanding Dzongkha.

Descriptive Analysis – Measure of Central Tendency

Table 241: Results of the SIM teachers' rating of SIM learning in understanding Dzongkha

```
. tabulate q13

        q13 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |         54        8.10        8.10
          2 |        309       46.33       54.42
          3 |        271       40.63       95.05
          4 |         33        4.95      100.00
------------+-----------------------------------
      Total |        667      100.00
```

From the frequency table above, it shows that mode choice is 2, which is "ineffective." The total SIM teacher respondents of only 45.5% chose "effective" or "extremely effective" for SIM learning in understanding Dzongkha.

Table 242: SIM teachers' rating of SIM learning in understanding Dzongkha, by age group

```
. tabulate age_group q13

           |                    q13
 Age_Group |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
   (20-24) |         1          1          3          0 |         5
   (25-29) |        15         87         75         11 |       188
   (30-34) |        18         86         93          6 |       203
   (35-39) |        15         73         56         10 |       154
   (40-44) |         2         33         29          1 |        65
   (45-49) |         2         19         13          2 |        36
   (50-54) |         1          7          2          2 |        12
   (55-59) |         0          3          0          1 |         4
-----------+--------------------------------------------+----------
     Total |        54        309        271         33 |       667
```

Looking at teachers' rating of SIM learning in understanding Dzongkha, by age group, it shows that results are mixed. Majority of the age groups have the mode as 2, which is "ineffective". However, age groups 20-24 and 30-34 have mode as 3 which is "effective."

Table 243: SIM teachers' rating of SIM learning in understanding Dzongkha, by key stage

```
. tabulate key_stage q13

              |                    q13
    Key_Stage |         1          2          3          4 |     Total
--------------+--------------------------------------------+----------
  Key Stage I |         9         66         54          9 |       138
 Key Stage II |        10         75         72          8 |       165
Key Stage III |        13         40         25          1 |        79
 Key Stage IV |        12         70         72         10 |       164
  Key Stage V |        10         58         48          5 |       121
--------------+--------------------------------------------+----------
        Total |        54        309        271         33 |       667
```

Looking at teachers' rating of SIM learning in understanding Dzongkha by key stage, it shows that majority of the key stages have the mode as 2, which is "ineffective".

Table 244: SIM teachers' rating of SIM learning in understanding Dzongkha, by school type

```
. tabulate school q13

           |                         q13
    School |         1          2          3          4 |     Total
-----------+--------------------------------------------+----------
       ECR |         1          1          3          0 |         5
       HSS |        23        135        115         17 |       290
       LSS |         6         25         19          2 |        52
       MSS |        22         98         70          5 |       195
        PS |         2         50         64          9 |       125
-----------+--------------------------------------------+----------
     Total |        54        309        271         33 |       667
```

Looking at teachers' rating of SIM learning in understanding Dzongkha by school type, it shows the majority of the school types except ECR and PS have mode as 2, which is "ineffective." However, ECR and PS has mode as 3, which is "effective."

Table 245: Median of the SIM teachers' rating of SIM learning in understanding Dzongkha

```
. tabstat q13, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q13 |       667         2         1         4
------------------------------------------------------
```

The calculated sample median = 2, which is "ineffective." This means at least 50% of the SIM teacher respondents found SIM learning "ineffective" or "extremely ineffective" in understanding Dzongkha.

Descriptive Analysis - Measure of Dispersion

Table 246: SIM teachers' measure of consensus on SIM learning in understanding Dzongkha

```
. cns q13 , min(1) max(4)

Consensus Measure for q13
Cns(X) = .64480862
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for SIM learning in understanding Dzongkha, it is 0.6448.

Inferential Analysis – Statistical Significance Testing

Table 247: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q13 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      304     99317.5      111389
    Negative |      363    123460.5      111389
        Zero |        0           0           0
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -4078538.3
Adjustment for zeros          0
                     ----------
Adjusted variance      20705514

H0: q13 = 2.5
         z = -2.653
Prob > |z| = 0.0080
Exact prob = 0.0080
```

We have seen that the 45.5% of SIM teachers surveyed think that SIM learning was effective or extremely effective in understanding Dzongkha. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM teacher population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "ineffective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM teacher population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM teacher population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0080, which is significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -2.653, p = 0.0080. The negative z-score shows that the population median is below the hypothesized median of 2.5.

Effect Size

The test statistic is Z = -2.653 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -2.653/SQRT(667) = -0.10. Ignoring negative sign, this, according to Bartz (1999), is a very low effect size.

Gender difference in SIM teachers' perception of SIM learning in understanding Dzongkha

Table 248: Results of Two-Sample Wilcoxon Rank-sum (Mann-Whitney) Test

```
. ranksum q13, by(gender) exact

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

      gender |      Obs    Rank sum    Expected
-------------+---------------------------------
      Female |      267       87043       89178
        Male |      400      135735      133600
-------------+---------------------------------
    Combined |      667      222778      222778

Unadjusted variance  5945200.00
Adjustment for ties  -993714.81
                     ----------
Adjusted variance    4951485.19

H0: q13(gender==Female) = q13(gender==Male)
         z = -0.959
Prob > |z| = 0.3373
Exact prob = 0.3380
```

There is no evidence for statistically significant difference between female teachers and male teachers (p-value = 0.3380 > alpha = 0.05) on perception of SIM learning in understanding Dzongkha, which means female teachers and male teachers rated similarly on SIM learning effectiveness in understanding Dzongkha.

Evidence on SIM Teachers' Perception of SIM Learning in Understanding Dzongkha

There is statistically significant evidence (p = 0.0080) that only minority 45.5% of SIM teachers, both female teachers and male teachers, found SIM learning effective in understanding Dzongkha. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -2.653, p = 0.0080, with a very low effect size (*r* = 0.10).

Advantages and Disadvantages of SIM Learning

Analyzing SIM Teachers' Perception of Advantages of SIM Learning

The Ministry of Education was interested to know what SIM teachers found as advantages of SIM learning. To investigate this, ***Figure 46*** shows the results of SIM teachers' perception of advantages of SIM learning.

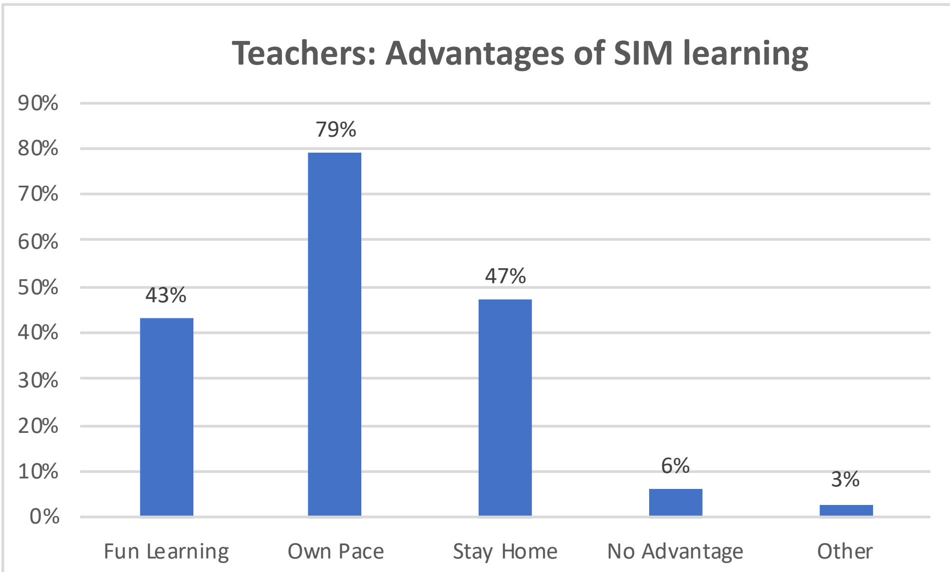


Figure 46: *Results of "What are the advantages of SIM-learning?"*

As shown in ***Figure 46,*** the SIM teachers found "Learning on your own pace" (79%) as the main advantage of SIM learning, followed by "Ability to stay at home" (47%) and "Self-learning is fun" (43%).

Inferential Analysis – Statistical Significance Testing through Cochran's Q Test

To test if the differences between advantages of SIM learning are significantly different we can use a Cochran's Q test.

Table 249: Results of Cochran's Q Test on Advantages of SIM Learning

```
. cochran q5_1 q5_2 q5_3 q5_4 q5_5, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
        q5_1 |   .4302849        287
        q5_2 |   .7916042        528
        q5_3 |   .4662669        311
        q5_4 |   .0554723         37
        q5_5 |   .0269865         18
------------------------------------

Number of obs        =        667
Cochran's chi2(4)    =   1073.172
Prob > chi2          =     0.0000
```

We have seen that the 79% of SIM teachers surveyed think that the main advantage of SIM learning was "Learning on your own pace," followed by "Ability to stay at home" (47%) and "Self-learning is fun" (43%). However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether there are differences between the proportions among the five options of advantages of SIM learning.

Ho: Our null hypothesis is that there are no differences between the proportions among the five options of advantages of SIM learning.

Ha: Our alternative hypothesis is that there are statistically significant differences between the proportions among the five options of advantages of SIM learning.

Cochran's Q test would show us how likely to have result as in our survey sample or even higher, if in the population there would be no differences. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that most likely in the population each option is not chosen equally often. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of advantages of SIM learning, χ^2 $(4, N = 667) = 1073.172$, $p = 0.0000$.

Post-hoc test

Since there are statistically significant differences in proportions of advantages of SIM learning, we would like to know whether there is statistically significant difference between "Learning on your own pace" (79%) and "Ability to stay at home" (47%) through pairwise comparisons as these two options were majority of the SIM teachers' choices on advantages of SIM learning. We will use Cochran's test for pairs.

Table 250: Results of Cochran's Q post-hoc test

```
. cochran q5_1 q5_2, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
        q5_1 |   .4302849        287
        q5_2 |   .7916042        528
------------------------------------

Number of obs        =        667
Cochran's chi2(1)    =   162.6919
Prob > chi2          =     0.0000
Exact p              =     0.0000
```

A pairwise post-hoc Cochran's Q test was statistically significant for "Learning at your own pace" vs. "Ability to stay at home", χ^2 (1, N = 667) = 162.6919, p = 0.0000. Therefore, the number one advantage of SIM learning for SIM teachers was "Learning at your own pace." The effect size between them $\eta^{2\,=}$ 162.6919/667 = 0.24.

Effect Size

The test statistic is Q = 1073.172, our sample size for SIM teachers is 667 and we have five options (variables) for advantages of SIM learning. Therefore, the effect size for this can be calculated by eta-squared (η^2) (Serlin, Carr, & Marascuilo, 1982).

η^2 = 1073.172/((5-1)x667) = 0.40, which is a large effect size.

Evidence on SIM Teachers' Perception on Advantages of SIM Learning

There is statistically significant evidence (p = 0.0000) that the majority of SIM teachers found "Learning at your own pace" as the main advantage of SIM learning, followed by "Ability to stay at home". In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of advantages of SIM learning, χ^2(4, N = 667) = 1073.172, p =0.0000, with a large effect size (η^2 = 0.40). A pairwise post-hoc Cochran test was also significant for "Learning at your own pace" vs. "Ability to stay at home" (p = .0000) with a moderate difference (η^2 = 0.24).

Analyzing SIM Teachers' Perception on Disadvantages of SIM Learning

The Ministry of Education was interested to know what SIM teachers found as disadvantages of SIM learning. To investigate this, ***Figure 47*** shows the results of SIM teachers' perception of disadvantages of SIM learning.

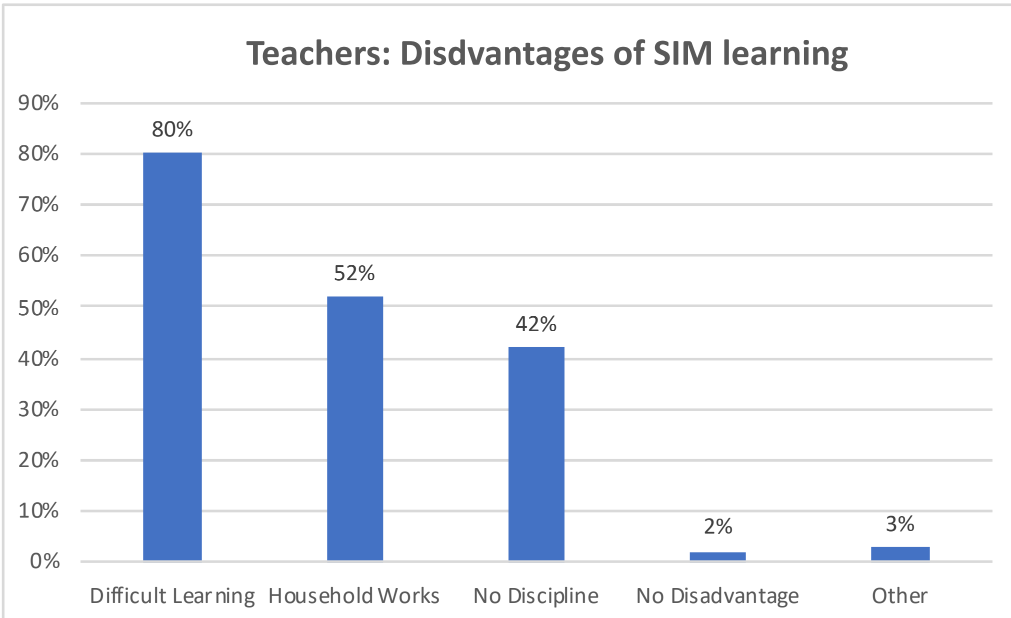


Figure 47: *Results of "What are the disadvantages of SIM-learning?"*

As shown in ***Figure 47,*** the SIM teachers found "Self-learning is difficult" (80%) as the main disadvantage of SIM learning, followed by "Household works at home" (52%) and "No self-discipline" (42%).

Inferential Analysis – Statistical Significance Testing through Cochran's Q Test

To test if the differences between disadvantages of SIM learning are significantly different we can use a Cochran's Q test.

Table 251: Results of Cochran's Q Test on Disadvantages of SIM Learning

```
. cochran q6_1 q6_2 q6_3 q6_4 q6_5, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+----------------------
        q6_1 |   .7946027        530
        q6_2 |   .5187406        346
        q6_3 |   .4182909        279
        q6_4 |   .0209895         14
        q6_5 |   .0344828         23
------------------------------------

Number of obs       =        667
Cochran's chi2(4)   =   1164.234
Prob > chi2         =     0.0000
```

We have seen that the 80% of SIM teachers surveyed think that the main disadvantage of SIM learning was "Self-learning is difficult," followed by "Household works at home" (52%) and "No self-discipline" (42%). However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether there are differences between the proportions among the five options of disadvantages of SIM learning.

Ho: Our null hypothesis is that there are no differences between the proportions among the five options of disadvantages of SIM learning.

Ha: Our alternative hypothesis is that there are statistically significant differences between the proportions among the five options of disadvantages of SIM learning.

Cochran's Q test would show us how likely to have result as in our survey sample or even higher, if in the population there would be no differences. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that most likely in the population each option is not chosen equally often. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of disadvantages of SIM learning, χ^2 $(4, N = 667) = 1164.234$, $p = 0.0000$.

Post-hoc test

Since there are statistically significant differences in proportions of disadvantages of SIM learning, we would like to know whether there is statistically significant difference between

“Self-learning is difficult” (80%) and “Household works at home” (52%) through pairwise comparisons as these two options are most selected of the SIM teachers’ choices on disadvantages of SIM learning. We will use Cochran’s test for pairs.

Table 252: Results of Cochran’s Q post-hoc test

```
. cochran q6_1 q6_2, detail

Test for equality of proportions of nonzero
outcomes in matched samples (Cochran's Q):

    Variable | Proportion      Count
-------------+---------------------
        q6_1 |   .7946027        530
        q6_2 |   .5187406        346
-----------------------------------

Number of obs       =        667
Cochran's chi2(1)   =   89.56614
Prob > chi2         =     0.0000
Exact p             =     0.0000
```

A pairwise post-hoc Cochran’s Q test was statistically significant for “Self-learning is difficult” vs. “Household works at home”, χ^2 (1, N = 667) = 89.56614, p = 0.0000. Therefore, the number one disadvantage of SIM learning for SIM teachers was “Self-learning is difficult.” The effect size between them $\eta^2 =$ 89.56614/667 = 0.13, which is a moderate effect size.

Effect Size

The test statistic is Q = 1164.234, our sample size for SIM teachers is 667 and we have five options (variables) for disadvantages of SIM learning. Therefore, the effect size for this can be calculated by eta-squared (η^2) (Serlin, Carr, & Marascuilo, 1982).

η^2 = 1164.234/((5-1)x667) = 0.44, which is a large effect size.

Evidence on SIM Teachers’ Perception of Disadvantages of SIM Learning

There is statistically significant evidence (p = 0.0000) that the majority of SIM teachers found “Self-learning is difficult” as the main disadvantage of SIM learning. In particular, Cochran's Q test indicated that there are differences between the proportions among the five options of disadvantages of SIM learning, χ^2(4, N = 667) = 1164.234, p =0.0000, with a large effect size (η^2 = 0.44). A pairwise post-hoc Cochran test was also significant for “Self-learning is difficult” vs. “Household works at home” (p = .0000) with a moderate effect size (η^2 = 0.13).

Effect of Household Chores on SIM Learning

Effect of Household Chores on SIM Learning: Is "Household works at home" a statistically significant disadvantage for the majority of the SIM students in the perception of SIM teachers?

As a social norm perception, usually people think having to do household works or chores at home is a disadvantage for studying at home, especially for adolescent girls during the COVID-19 pandemic. In this SIM program assessment study, we surveyed and tested this perception too. We found 52% of the SIM teachers surveyed selected "Household works at home" as a disadvantage for their students during SIM learning. We need to test whether the majority of the SIM teachers in the population would select "Household works at home" as a disadvantage or not.

Ho: Our null hypothesis is that the percentage of the SIM teachers who selected "Household works at home" as a disadvantage is 50%.

Ha: Our alternative hypothesis is the percentage of the SIM teachers who selected "Household works at home" as a disadvantage is greater than 50%.

Table 253: Results of One Sample Binomial Test on Household Works

```
. bitest q6_2 = 0.50

Binomial probability test

    Variable |            N   Observed k   Expected k   Assumed p
Observed p
-------------+-------------------------------------------------------------
----
        q6_2 |          667          346        333.5     0.50000
0.51874

  Pr(k >= 346)            = 0.176375  (one-sided test)
  Pr(k <= 346)            = 0.842967  (one-sided test)
  Pr(k <= 321 or k >= 346) = 0.352750  (two-sided test)
```

One-sided binomial test indicated that the percentage of SIM teachers who selected "Household works at home" as a disadvantage (N_{hw} = 346, 52%), was not statistically significantly different from the population hypothesized value of 50%, p = 0.176375 (which greater than alpha = 0.05). Therefore, there is no sufficient evidence that "Household works at home" affected the majority of SIM students during SIM learning even in the perception of SIM teachers.

Gender Difference on Effect of Household Chores on SIM Learning

Table 254: Results of Two-Sample Test of Proportions on Household Works, by Gender

```
. prtest q6_2, by(gender)

Two-sample test of proportions           Female: Number of obs =      267
                                           Male: Number of obs =      400
------------------------------------------------------------------------------
 Group |       Mean   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
Female |   .5205993   .0305735                      .4606762    .5805223
  Male |      .5175   .0249847                      .4685309    .5664691
-------------+----------------------------------------------------------------
  diff |   .0030993   .0394839                     -.0742877    .0804862
             |  under H0:   .0394859     0.08   0.937
------------------------------------------------------------------------------
  diff = prop(Female) - prop(Male)                               z =   0.0785
    H0: diff = 0

  Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
Pr(Z < z) = 0.5313         Pr(|Z| > |z|) = 0.9374          Pr(Z > z) = 0.4687
```

Since our SIM survey sample is large enough (N=667) to assume normal distribution, we applied two-sample test of proportions to test whether "Household works at home" affected girls more than boys during SIM learning in times of COVID-19 pandemic. We found that there is no statistically significant evidence that girls were affected more than boys by "Household works at home" during the SIM learning, $z = 0.0785$, $p = 0.4687$ (which is greater than alpha = 0.05). Therefore, "Household works at home" was not statistically significant disadvantage for the in the perception of SIM teachers, both female teachers and male teachers, during SIM learning.

Help Given for SIM Learning

Analyzing SIM Teachers' Perception of Help Given for SIM Learning

The Ministry of Education was interested to know if SIM teachers gave help during SIM learning. To investigate this, ***Figure 48*** shows the results of SIM teachers' perception on help given during SIM learning.

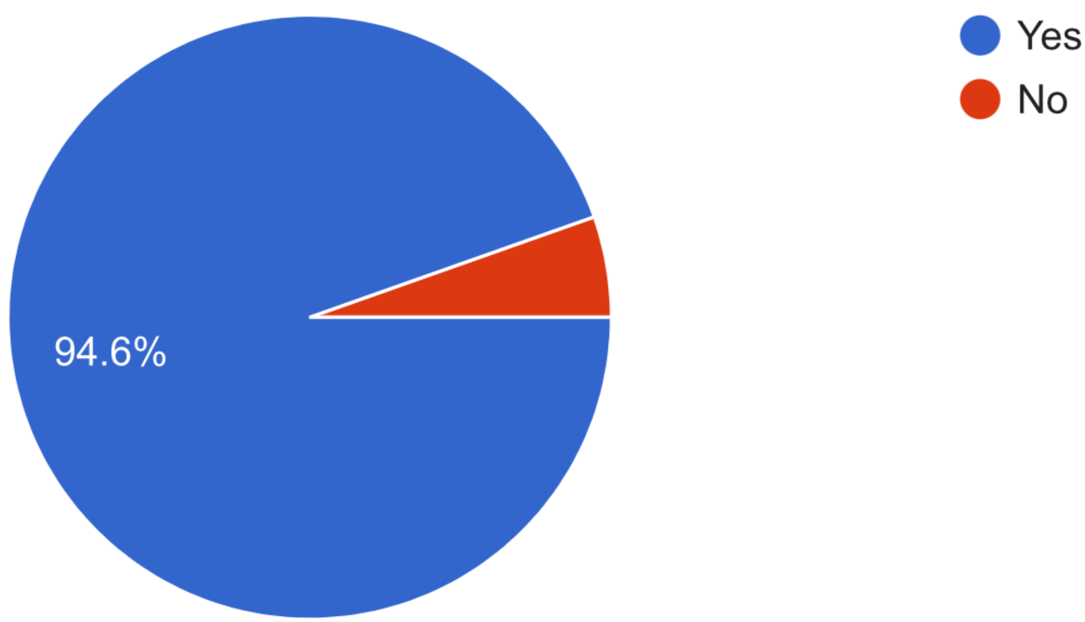


Figure 48: *Results of "Did you give help to anyone to understand SIM lessons?"*

As shown in ***Figure 48,*** the 94.6% of SIM teachers said they gave help to someone to understand SIM lessons.

Inferential Analysis – Statistical Significance Testing

Table 255: Results of Binomial Test on Help Given for SIM lessons

```
. bitest q28a = 0.92

Binomial probability test

 Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q28a |        667          631       613.64     0.92000      0.94603

  Pr(k >= 631)            = 0.005874  (one-sided test)
  Pr(k <= 631)            = 0.996322  (one-sided test)
  Pr(k <= 595 or k >= 631) = 0.012255  (two-sided test)
```

A binomial test indicated that the percentage of SIM teachers who gave help for SIM lessons

(N_{help} = 631, 94.6%), was statistically significantly greater than the population hypothesized value of 92%, p = 0.005874.

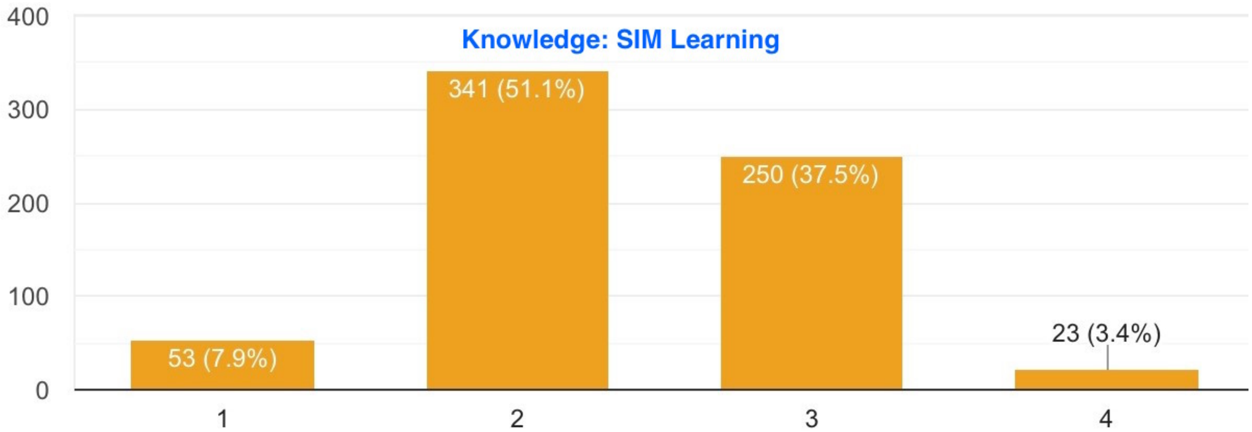
Evidence on SIM Teachers' Help Given for SIM Lessons

There is statistically significant evidence (p = 0.005874) that at least 92% of SIM teachers gave help for SIM lessons. In other words, a binomial test indicated that the percentage of SIM teachers who gave help for SIM lessons (N_{help} = 631, 94.6%) was statistically significantly greater than the population hypothesized value of 92%, p = 0.005874.

Comparison between SIM Learning and Classroom Learning

Effectiveness of SIM learning vs Classroom Learning in increasing knowledge

The Ministry of Education was interested to know how SIM teachers found SIM learning in increasing knowledge in comparison to classroom learning. To investigate this, ***Figure 49*** shows the results of SIM teachers' perception on increasing knowledge during SIM learning in comparison to classroom learning.



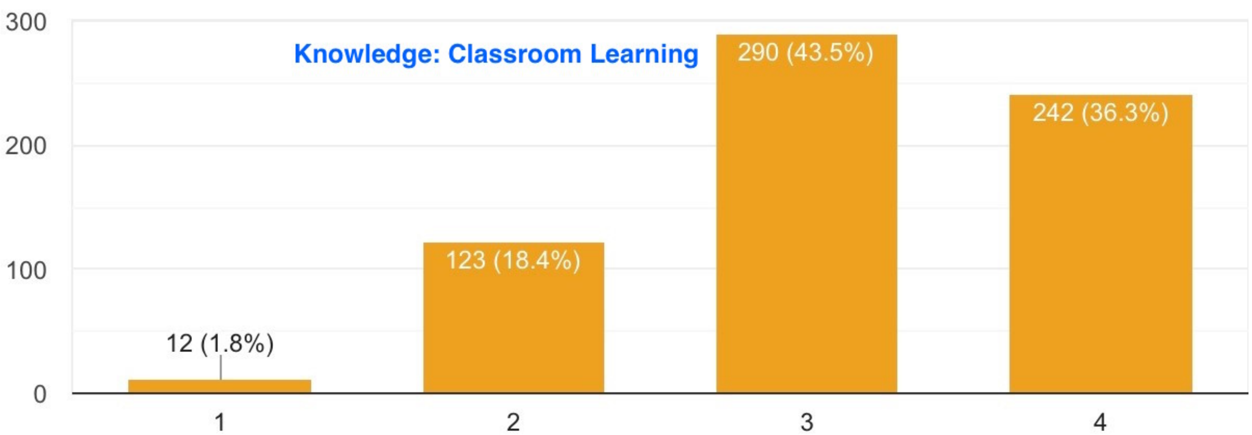


Figure 49*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of increasing knowledge" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 49*** the 40.9% (SIM) vs 79.8% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in increasing knowledge.

Inferential Analysis – Statistical Significance Testing

Table 256: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q7 = q14, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       46       17733       98450
    Negative |      394      179167       98450
        Zero |      227       25878       25878
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -543913.13
Adjustment for zeros  -981207.5
                     ----------
Adjusted variance      23258932

H0: q7 = q14
         z = -16.737
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that only 40.9% of SIM teachers surveyed think that SIM learning was effective or extremely effective in increasing knowledge. Comparing it with classroom learning, about 79.8% of the same group of SIM teachers surveyed think that classroom learning was effective or extremely effective in increasing knowledge. Classroom learning was more effective in increasing knowledge. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for increasing knowledge is significantly different from the true median of classroom learning in increasing knowledge in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of increasing knowledge.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of increasing knowledge.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of increasing knowledge. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -16.737, p = 0.0000.

The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -16.737 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -16.737/SQRT(667) = -0.65 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Increasing Knowledge

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in increasing knowledge. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of increasing knowledge, Z = -16.737, p = 0.0000, with a strong effect size or difference (*r* = 0.65).

Effectiveness of SIM learning vs Classroom Learning in increasing skills

The Ministry of Education was interested to know how SIM teachers found SIM learning in increasing skills in comparison to classroom learning. To investigate this, ***Figure 50*** shows the results of SIM teachers' perception on increasing skills during SIM learning in comparison to classroom learning.

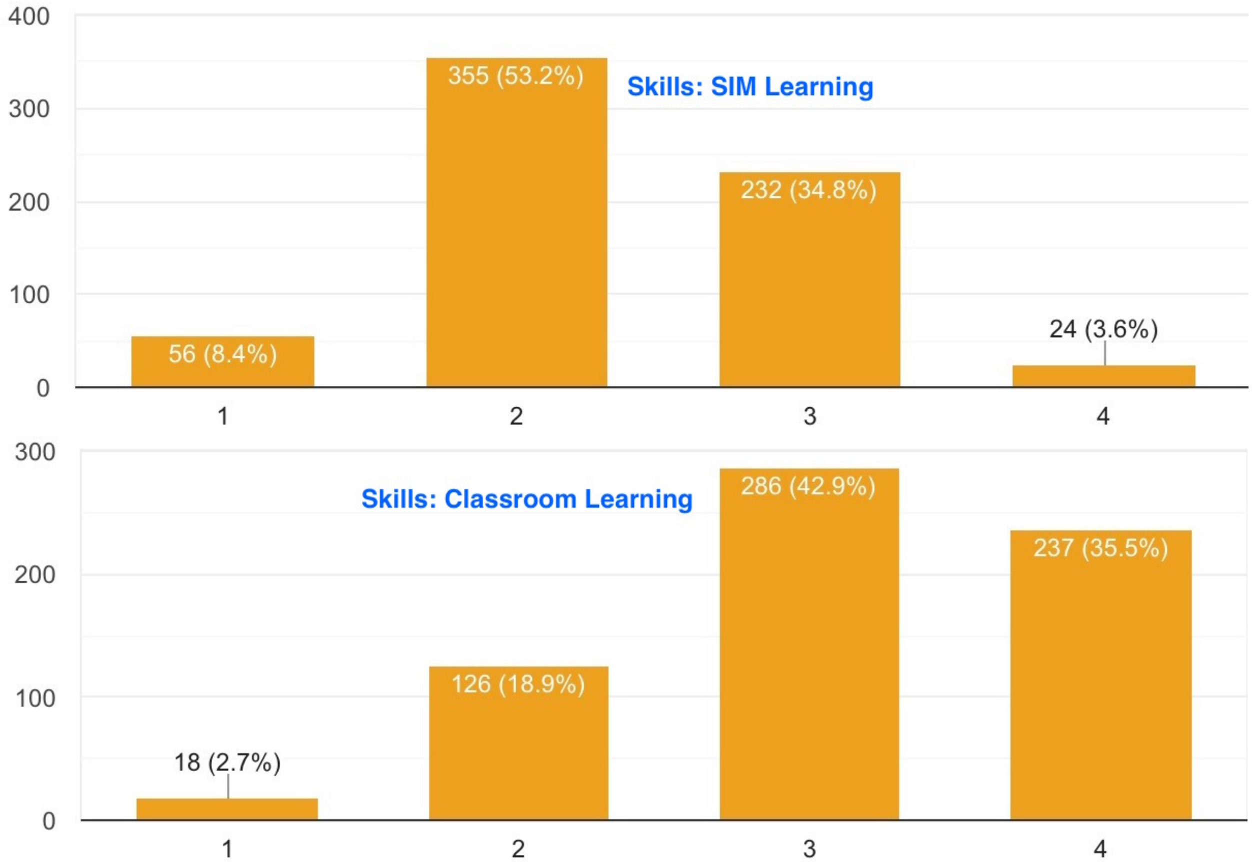


Figure 50*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of increasing skills" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 50*** the 38.4% (SIM) vs 78.4% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in increasing skills.

Inferential Analysis – Statistical Significance Testing

Table 257: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = q15, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       49       18589     98106.5
    Negative |      388      177624     98106.5
        Zero |      230       26565       26565
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties   -507374.5
Adjustment for zeros -1020538.7
                     ----------
Adjusted variance      23256139

H0: q8 = q15
         z = -16.489
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 38.4% of SIM teachers surveyed think that SIM learning was effective or extremely effective in increasing skills. Comparing it with classroom learning, about 78.4% of the same group of SIM teachers surveyed also think that classroom learning was effective or extremely effective in increasing skills. Classroom learning was more effective in increasing skills. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for increasing skills is significantly different from the true median of classroom learning in increasing skills in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of increasing skills.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of increasing skills.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of increasing skills. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -16.489, p = 0.0000. The negative

z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -16.489 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -16.489/SQRT(667) = -0.64 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Increasing Skills

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in increasing skills. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of increasing skills, Z = -16.489, p = 0.0000, with a strong effect size or strong difference (r = 0.64).

Effectiveness of SIM learning vs Classroom Learning in imparting values

The Ministry of Education was interested to know how SIM teachers found SIM learning in imparting values. To investigate this, ***Figure 51*** shows the results of SIM teachers' perception on imparting values during SIM learning in comparison to classroom learning.

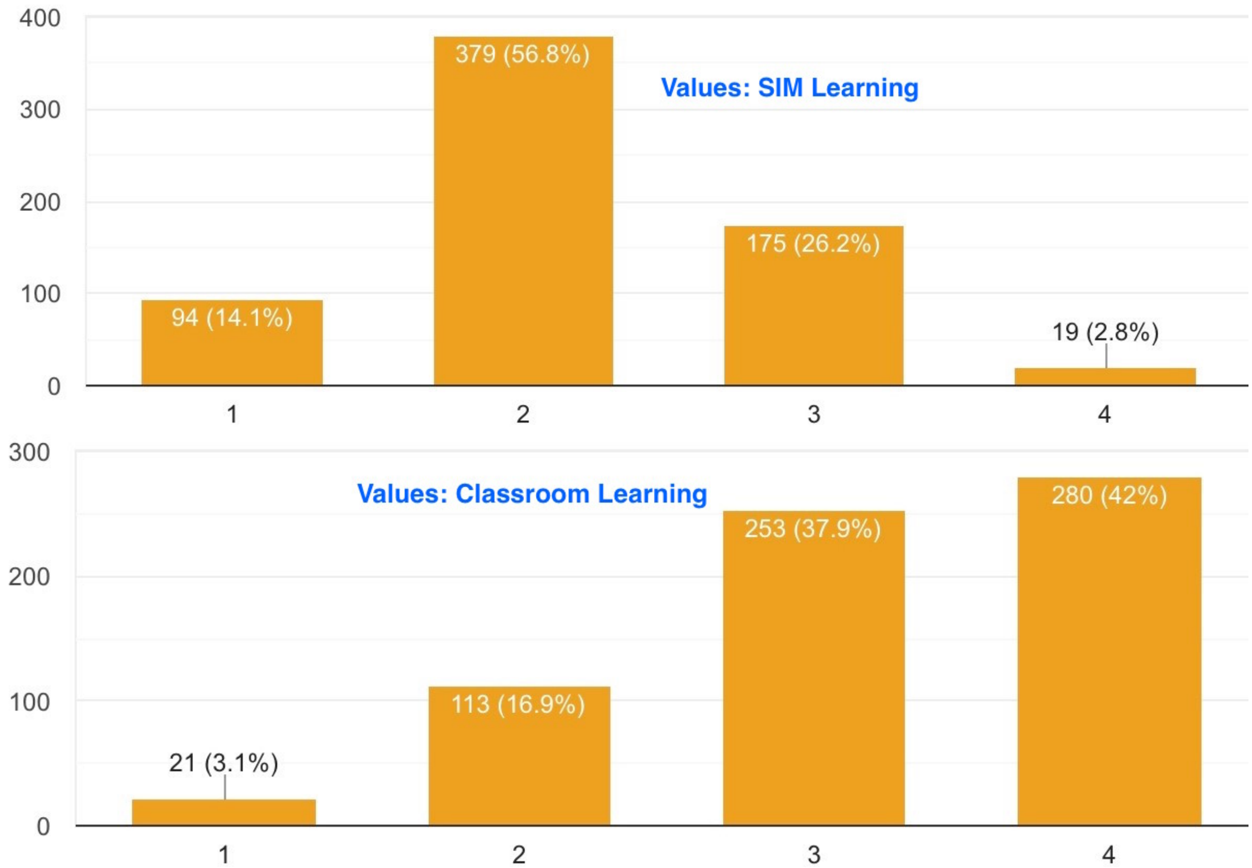


Figure 51*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of imparting values" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 51*** the 29.0% (SIM) vs 79.9% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in imparting values.

Inferential Analysis – Statistical Significance Testing

Table 258: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q9 = q16, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       46       14896      102600
    Negative |      434      190304      102600
        Zero |      187       17578       17578
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties   -429512.5
Adjustment for zeros  -549312.5
                     ----------
Adjusted variance      23805228

H0: q9 = q16
         z = -17.976
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 29.0% of SIM teachers surveyed think that SIM learning was effective or extremely effective in imparting values. Comparing it with classroom learning, about 79.9% of the same group of SIM teachers surveyed also think that classroom learning was effective or extremely effective in imparting values. Classroom learning was more effective in imparting values. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for imparting values is significantly different from the true median of classroom learning in imparting values in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of imparting values.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of imparting values.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of imparting values. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -17.976, p = 0.0000. The negative

z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -17.976 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -17.976/SQRT(667) = -0.70 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Imparting Values

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in imparting values. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of imparting values, Z = -17.976, p = 0.0000, with a strong effect size or strong difference (r = 0.70).

Effectiveness of SIM learning vs Classroom Learning in improving attitudes

The Ministry of Education was interested to know how SIM teachers found SIM learning in improving attitudes. To investigate this, ***Figure 52*** shows the results of SIM teachers' perception on improving attitudes during SIM learning in comparison to classroom learning.

Figure 52*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of improving attitudes" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 52*** the 23.1% (SIM) vs 77.9% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in improving attitudes.

Inferential Analysis – Statistical Significance Testing

Table 259: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q10 = q17, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       33     10147.5    103423.5
    Negative |      456    196699.5    103423.5
        Zero |      178       15931       15931
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -461764.63
Adjustment for zeros -473947.25
                     ----------
Adjusted variance      23848341

H0: q10 = q17
         z = -19.100
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 23.1% of SIM teachers surveyed think that SIM learning was effective or extremely effective in improving attitudes. Comparing it with classroom learning, 77.9% of the same group of SIM teachers surveyed also think that classroom learning was effective or extremely effective in improving attitudes. Classroom learning was more effective in improving attitudes. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for improving attitudes is significantly different from the true median of classroom learning in improving attitudes in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of improving attitudes.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of improving attitudes.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of improving attitudes. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -19.100, p = 0.0000. The negative

z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -19.100 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -19.100/SQRT(667) = -0.74 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Improving Attitudes

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in improving attitudes. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of improving attitudes, Z = -19.100, p = 0.0000, with a strong effect size or strong difference (*r* = 0.74).

Effectiveness of SIM learning vs Classroom Learning in understanding English

The Ministry of Education was interested to know how SIM teachers found SIM learning in understanding English. To investigate this, ***Figure 53*** shows the results of SIM teachers' perception on understanding English during SIM learning in comparison to classroom learning.

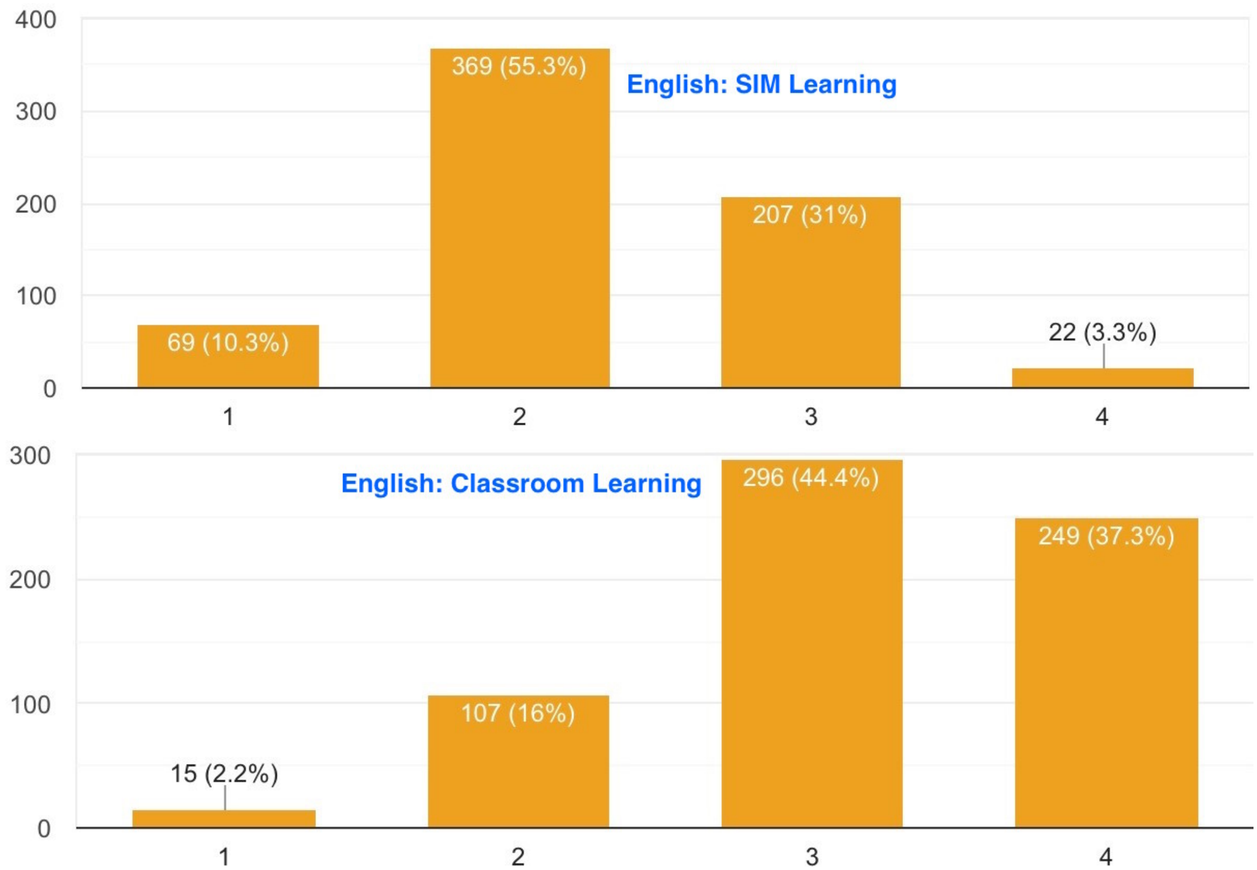


Figure 53*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of understanding English" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 53*** the 34.3% (SIM) vs 81.7% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in understanding English.

Inferential Analysis – Statistical Significance Testing

Table 260: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q11 = q18, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       32     11471.5       99234
    Negative |      415    186996.5       99234
        Zero |      220       24310       24310
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties   -452824.5
Adjustment for zeros  -893392.5
                      ----------
Adjusted variance      23437836

H0: q11 = q18
         z = -18.128
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 34.3% of SIM teachers surveyed think that SIM learning was effective or extremely effective in understanding English. Comparing it with classroom learning, 81.7% of the same group of SIM teachers surveyed also think that classroom learning was effective or extremely effective in understanding English. Classroom learning was more effective in understanding English. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for understanding English is significantly different from the true median of classroom learning in understanding English in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of understanding English.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of understanding English.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of understanding English. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -18.128, p = 0.0000.

The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -18.128 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -18.128/SQRT(667) = -0.70 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Understanding English

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in understanding English. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of understanding English, Z = -18.128, p = 0.0000, with a strong effect size or strong difference (r = 0.70).

Effectiveness of SIM learning vs Classroom Learning in understanding Maths

The Ministry of Education was interested to know how SIM teachers found SIM learning in understanding Mathematics. To investigate this, ***Figure 54*** shows the results of SIM teachers' perception on understanding Mathematics during SIM learning in comparison to classroom learning.

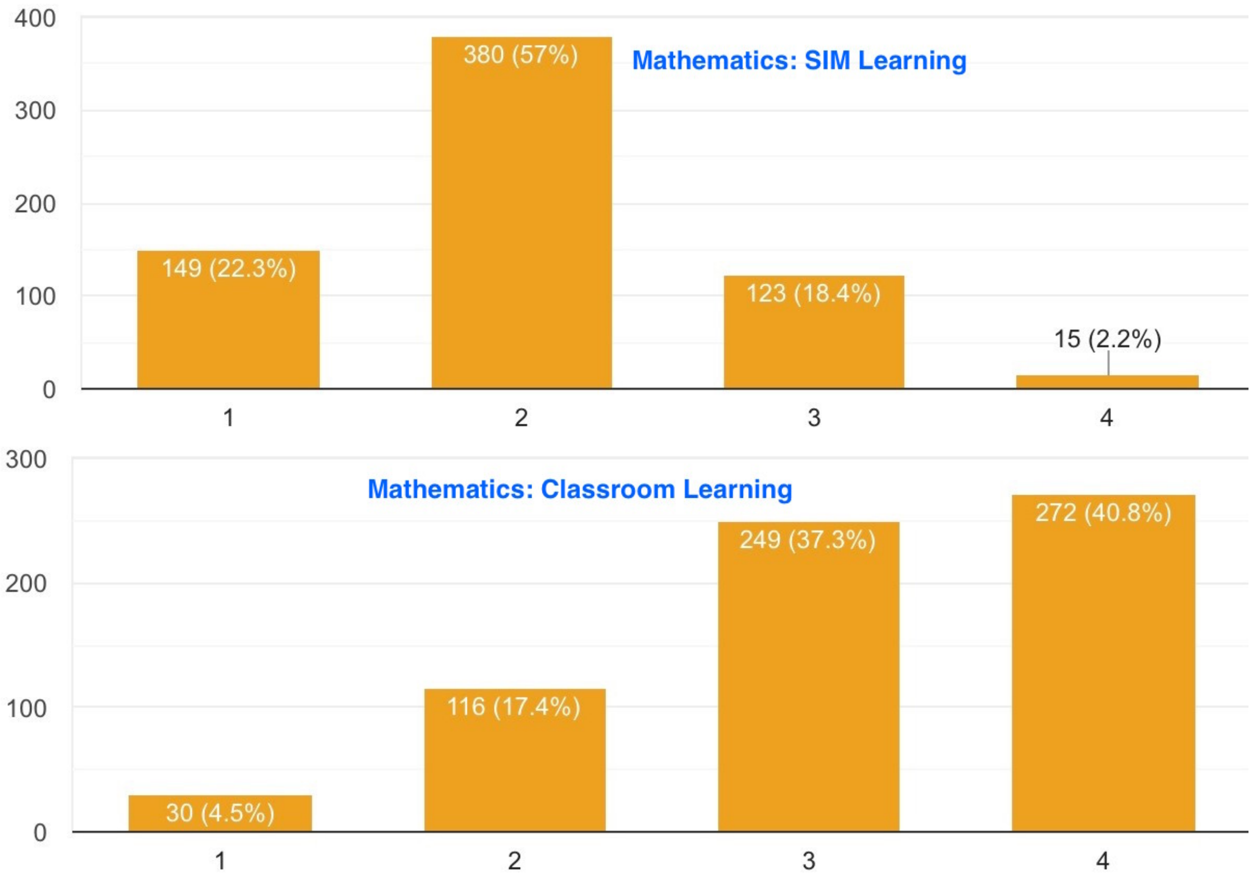


Figure 54*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of understanding Mathematics" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 54*** the 20.6% (SIM) vs 78.1% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in understanding Mathematics.

Inferential Analysis – Statistical Significance Testing

Table 261: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q12 = q19, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       35     10485.5      104036
    Negative |      461    197586.5      104036
        Zero |      171       14706       14706
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties   -413407.5
Adjustment for zeros  -420346.5
                     ----------
Adjusted variance      23950299

H0: q12 = q19
         z = -19.116
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 20.6% of SIM teachers surveyed think that SIM learning was effective or extremely effective in understanding Mathematics. Comparing it with classroom learning, 78.1% of the same group of SIM teachers surveyed also think that classroom learning was effective or extremely effective in understanding Mathematics. Classroom learning was more effective in understanding Mathematics. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for understanding Mathematics is significantly different from the true median of classroom learning in understanding Mathematics in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of understanding Mathematics.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of understanding Mathematics.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of understanding Mathematics. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -19.116, p = 0.0000. The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -19.116 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -19.116/SQRT(667) = -0.74 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Understanding Maths

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in understanding Mathematics. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of understanding Mathematics, Z = -19.116, p = 0.0000, with a strong effect size or difference (*r* = 0.74).

Effectiveness of SIM learning vs Classroom Learning in understanding Dzongkha

The Ministry of Education was interested to know how SIM teachers found SIM learning in understanding Dzongkha. To investigate this, ***Figure 55*** shows the results of SIM teachers' perception on understanding Dzongkha during SIM learning in comparison to classroom learning.

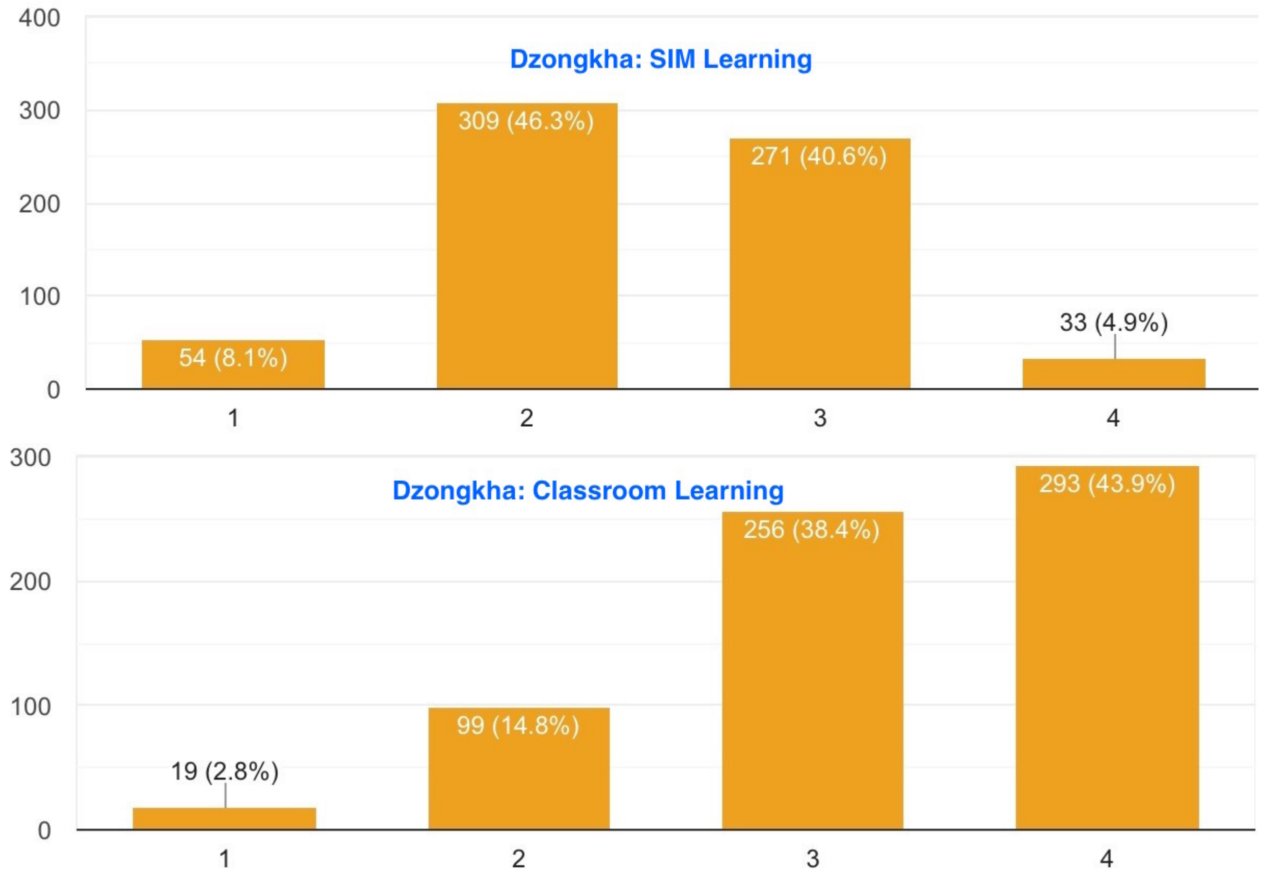


Figure 55*: Results of "Rate the effectiveness of SIM-learning vs Classroom-learning in terms of understanding Dzongkha" where 1 = Extremely ineffective, 2 = Ineffective, 3 = Effective, and 4 = Extremely effective*

As can be seen in ***Figure 55*** the 45.5% (SIM) vs 82.3% (Classroom) of the SIM teacher respondents rated "effective" or "extremely effective" in understanding Dzongkha.

Inferential Analysis – Statistical Significance Testing

Table 262: Results of Two-Sample Wilcoxon Signed Rank Test

```
. signrank q13 = q20, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       43     16208.5       97991
    Negative |      393    179773.5       97991
        Zero |      231       26796       26796
-------------+---------------------------------
         All |      667      222778      222778

Unadjusted variance    24784053
Adjustment for ties  -469923.25
Adjustment for zeros   -1033879
                     ----------
Adjusted variance      23280250

H0: q13 = q20
         z = -16.950
Prob > |z| =  0.0000
Exact prob =  0.0000
```

We have seen that the 45.5% of SIM teachers surveyed think that SIM learning was effective or extremely effective in understanding Dzongkha. Comparing it with classroom learning, 82.3% of the same group of SIM teachers surveyed also think that classroom learning was effective or extremely effective in understanding Dzongkha. Classroom learning was more effective in understanding Dzongkha. However, this was based on our sample from the SIM survey. We need to test whether this would be true in the SIM teacher population too. In other words, we have to test whether the true median of SIM learning for understanding Dzongkha is significantly different from the true median of classroom learning in understanding Dzongkha in the population.

Ho: Our null hypothesis is that there is no difference between true median of SIM learning and true median of classroom learning in terms of understanding Dzongkha.

Ha: Our alternative hypothesis is that the SIM teacher population's true median of SIM learning is significantly different from true median of classroom learning in terms of understanding Dzongkha.

Two-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if there was no significant difference between the true median in the population for SIM learning and classroom learning in terms of understanding Dzongkha. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median for SIM learning is statistically significantly different from true population median for classroom learning. In short, our two-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = -16.950, p = 0.0000. The negative z-score shows that the population median for SIM learning is less than the population median for classroom learning.

Effect Size

The test statistic is Z = -16.950 and our sample size for SIM teachers is 667. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is -16.950/SQRT(667) = -0.66 (we can ignore the negative sign). This, according to Bartz (1999), is strong effect size or strong difference.

Evidence on SIM Learning vs Classroom Learning in Understanding Dzongkha

There is statistically significant evidence (p = 0.0000) that the SIM teachers found classroom learning more effective than SIM learning in understanding Dzongkha. In particular, two-sample Wilcoxon signed rank test indicated that the teachers tend to like classroom learning more than SIM learning in terms of understanding Dzongkha, Z = -16.950, p = 0.0000, with a strong effect size or strong difference (*r* = 0.66).

PART III: SIM PRINCIPALS

Demographic Characteristics of SIM Principal Respondents

The age characteristics of the SIM principal respondents are summarized in ***Table 263***. The age of the SIM principal respondents ranged from 28 to 65 years (M = 43.17, SD = 6.34).

Table 263: Results of age characteristics of SIM principal respondents

Variable	Obs	Mean	Std. dev.	Min	Max
age	123	43.17073	6.338417	28	65

Similarly, among the 123 SIM principal respondents, 121 (98.4%) were males and 2 (1.6%) were females as shown in ***Figure 56***. The low representation of female principals is the reflection of reality in the population through random sampling and not lack of data.

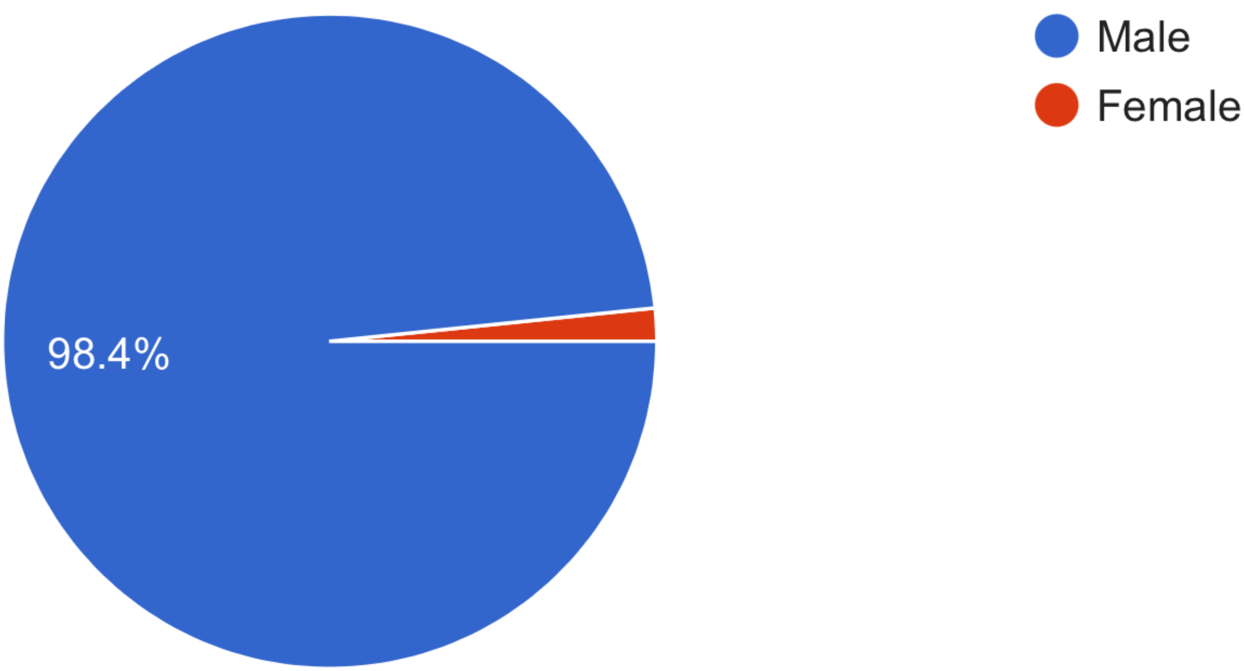


Figure 56: *Gender of SIM principal respondents*

Among the 123 SIM principal respondents, we got data representation from all types of schools such as HSS (18.7%), MSS (11.4%), LSS (7.3%), PS (57.7%), and ECR (4.9%) as shown in ***Figure 57***.

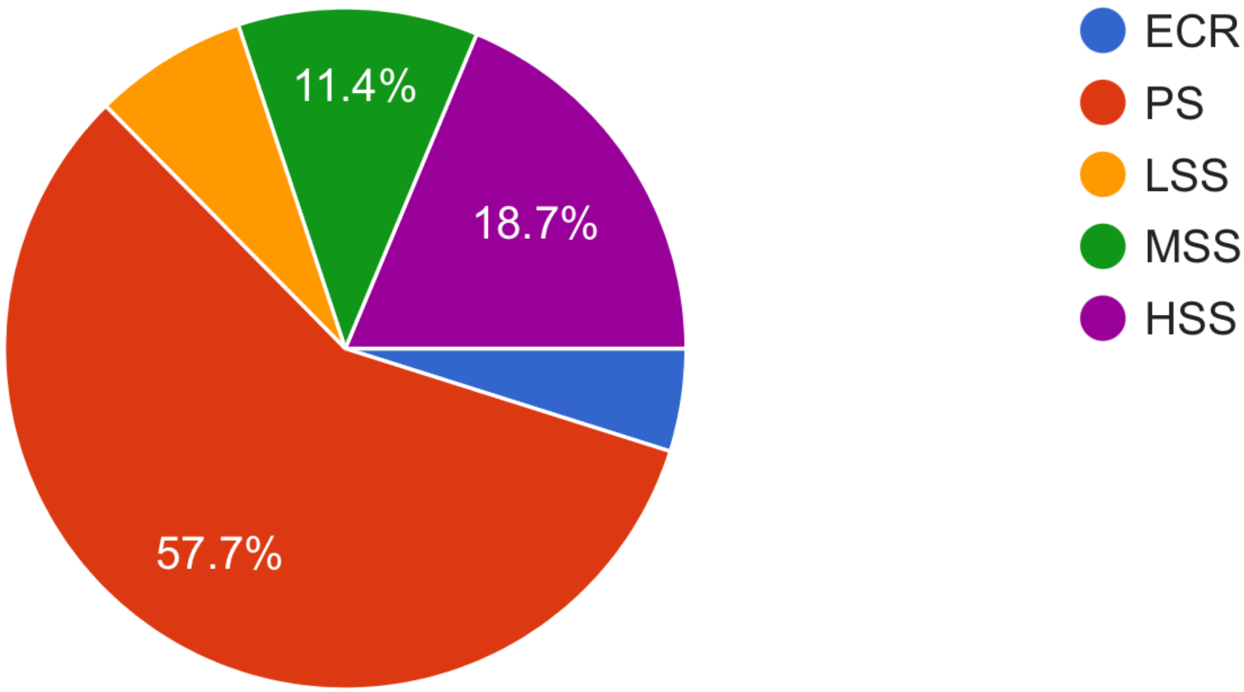


Figure 57*: School types of SIM principal respondents*

Effectiveness of SIM Programme

Analyzing Principals' Satisfaction Level of SIM

The Ministry of Education was interested to know satisfaction level of SIM programme, including principals' satisfaction level, during COVID-19 pandemic. To investigate this, ***Figure 58***, which is visualization of survey data, shows the results of satisfaction level of principals from the SIM survey.

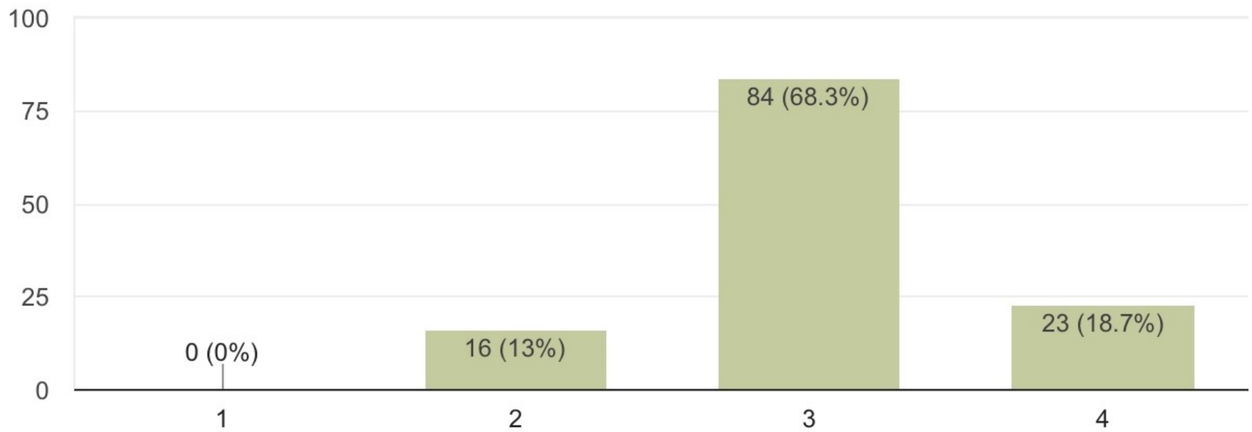


Figure 58*: Results of "Rate how satisfied are you with the current SIM" where 1 = Extremely dissatisfied, 2 = Dissatisfied, 3 = Satisfied, and 4 = Extremely satisfied*

As can be seen in ***Figure 58*** the 87.0% of the SIM principal respondents rated the SIM programme "satisfied" or "extremely satisfied."

Descriptive Analysis – Measure of Central Tendency

Table 264: Results of the SIM Principals' satisfaction level rating frequency distribution

```
. tabulate q12

        q12 |      Freq.     Percent        Cum.
------------+-----------------------------------
          2 |         16       13.01       13.01
          3 |         84       68.29       81.30
          4 |         23       18.70      100.00
------------+-----------------------------------
      Total |        123      100.00
```

From the frequency ***Table 264*** above, it shows that mode choice is 3, which is "satisfied." The total SIM principal respondents of 87.0% chose "satisfied" or "extremely satisfied."

Table 265: SIM Principals' satisfaction level rating frequency distribution, by school type

```
. tabulate school q12

           |               q12
    School |         2          3          4 |     Total
-----------+---------------------------------+----------
       ECR |         0          6          0 |         6
       HSS |         6         16          1 |        23
       LSS |         0          9          0 |         9
       MSS |         4          9          1 |        14
        PS |         6         44         21 |        71
-----------+---------------------------------+----------
     Total |        16         84         23 |       123
```

Looking at principals' satisfaction level of SIM survey data by school type, it shows that consistently in all school types, the mode is 3, which is "satisfied."

Table 266: Result of the SIM principals' satisfaction level rating median calculation

```
. tabstat q12, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q12 |       123         3         2         4
------------------------------------------------------
```

The calculated sample median = 3, which is "satisfied." This means at least 50% of the SIM principal respondents are in the "satisfied" or "extremely satisfied" category looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 267: Result of the SIM Principals' measure of consensus on satisfaction level

```
. cns q12 , min(1) max(4)

Consensus Measure for q12
Cns(X) = .79780574
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the satisfaction level of SIM principals, it is 0.7978.

Inferential Analysis – Statistical Significance Testing

Table 268: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q12 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      107        6818        3813
    Negative |       16         808        3813
        Zero |        0           0           0
-------------+---------------------------------
         All |      123        7626        7626

Unadjusted variance   156968.50
Adjustment for ties   -21084.25
Adjustment for zeros       0.00
                     ----------
Adjusted variance     135884.25

H0: q12 = 2.5
         z =  8.152
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 87.0% of SIM principals surveyed think that SIM programme was satisfactory. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM principal population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "dissatisfied" and 3 = "satisfied."

Ho: Our null hypothesis is that the SIM principal population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM principal population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.152, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 8.152 and our sample size for SIM Principals is 123. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 8.152/SQRT(123) = 0.74. This, according to Bartz (1999), is strong effect size.

Evidence on SIM Principals' Satisfaction Level

There is statistically significant evidence (p = 0.0000) that 87.0% of SIM principals are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.152, p = 0.0000, with a strong effect size (r = 0.74).

Analyzing Principals' Perception on Implementation of SIM

The Ministry of Education was interested to know how effectively implementation of SIM programme was carried out in the perception of principals. To investigate this, ***Figure 59***, which is visualization of survey data, shows the results of implementation effectiveness perception of principals from the SIM survey.

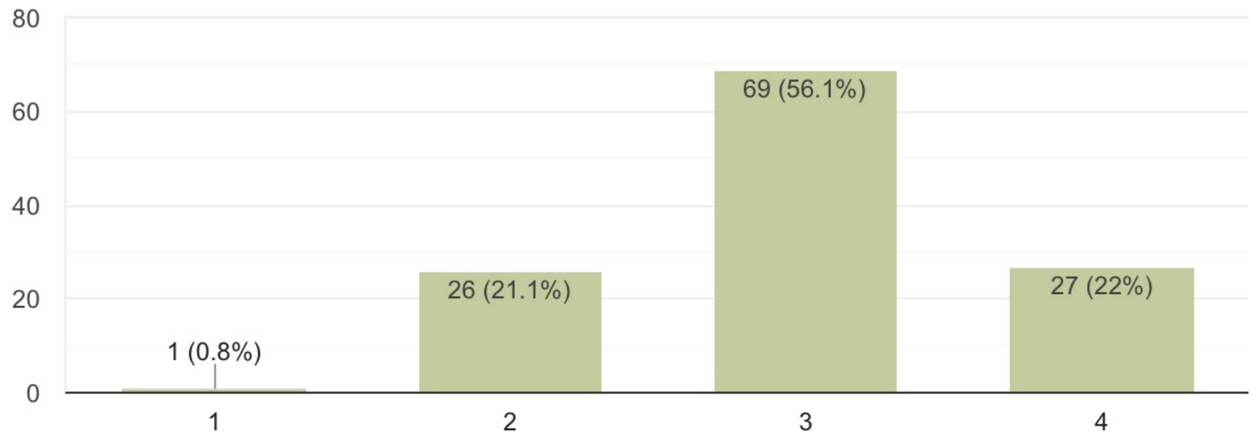


Figure 59: *Results of "Rate how effectively has the SIM been implemented" where 1 = Not effective, 2 = Slightly effective, 3 = Effective, and 4 = Very effective*

As can be seen in ***Figure 59*** the 78.1% of the SIM principal respondents rated that the SIM programme implementation was "effective" or "very effective."

Descriptive Analysis – Measure of Central Tendency

Table 269: Results of the SIM principals' SIM implementation rating frequency distribution

```
. tabulate q11

        q11 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          1        0.81        0.81
          2 |         26       21.14       21.95
          3 |         69       56.10       78.05
          4 |         27       21.95      100.00
------------+-----------------------------------
      Total |        123      100.00
```

From the frequency ***Table 269*** above, it shows that mode choice is 3, which is "effective." The total SIM principal respondents of 78.1% chose "effective" or "very effective" in their perception on implementation effectiveness of the SIM.

Table 270: SIM principals' SIM implementation rating frequency distribution, by school type

```
. tabulate school q11
```

School	q11 1	2	3	4	Total
ECR	0	0	4	2	6
HSS	0	11	12	0	23
LSS	0	2	6	1	9
MSS	0	5	8	1	14
PS	1	8	39	23	71
Total	1	26	69	27	123

Looking at principals' perception on implementation effectiveness of SIM by school type, it shows that consistently in all school types, the mode is 3, which is "effective."

Table 271: Result of the SIM principals' SIM implementation rating median calculation

```
. tabstat q11, stat(count p50 min max)
```

Variable	N	p50	Min	Max
q11	123	3	1	4

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM principal respondents believe that SIM implementation was "effective" or "very effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 272: Result of the SIM Principals' measure of consensus on SIM implementation

```
. cns q11 , min(1) max(4)

Consensus Measure for q11
Cns(X) = .732903
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the implementation effectiveness opinion of SIM principals, it is 0.7329.

Inferential Analysis – Statistical Significance Testing

Table 273: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q11 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       96      6268.5        3813
    Negative |       27      1357.5        3813
        Zero |        0           0           0
-------------+---------------------------------
         All |      123        7626        7626

Unadjusted variance   156968.50
Adjustment for ties   -18316.75
Adjustment for zeros       0.00
                     ----------
Adjusted variance     138651.75

H0: q11 = 2.5
         z =   6.594
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 78.1% of SIM principals surveyed think that SIM programme was effectively implemented. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM principal population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = “slightly effective” and 3 = “effective.”

Ho: Our null hypothesis is that the SIM principal population’s true median is 2.5.

Ha: Our alternative hypothesis is that the SIM principal population’s true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.594, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 6.594 and our sample size for SIM Principals is 123. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 6.594/SQRT(123) = 0.59. This, according to Bartz (1999), is moderate effect size.

Evidence on SIM Principals' Perception on Implementation Effectiveness of SIM

There is statistically significant evidence (p = 0.0000) that 78.1% of SIM principals believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.594, p = 0.0000, with a moderate effect size (*r* = 0.59).

Analyzing Principals' Perception on Usefulness of SIM

The Ministry of Education was interested to know how useful was SIM programme in the perception of principals. To investigate this, ***Figure 60***, which is visualization of survey data, shows the results of SIM usefulness perception of principals from the SIM survey.

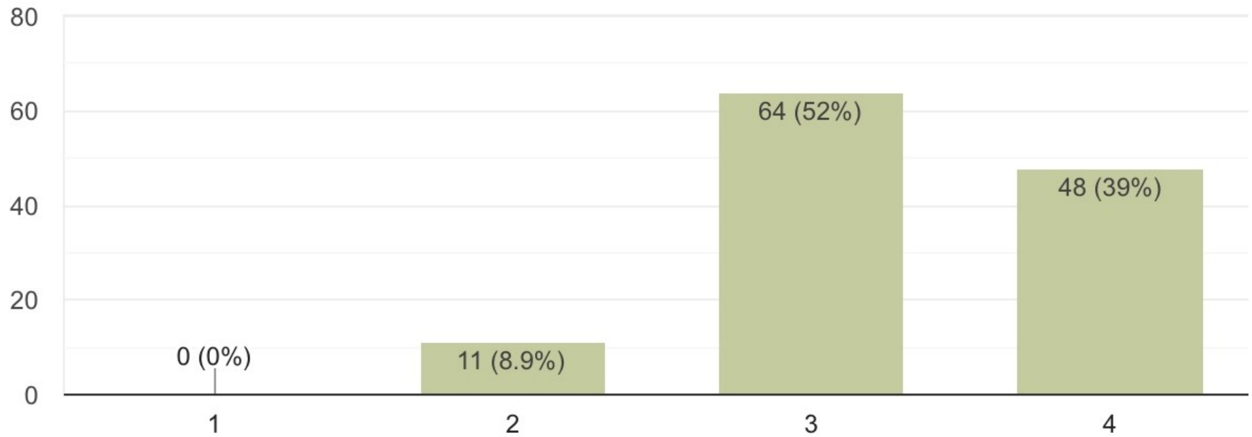


Figure 60: *Results of "Rate how useful was SIM" where 1 = Not useful, 2 = Slightly useful, 3 = Useful, and 4 = Very useful*

As can be seen in ***Figure 60*** the 91.0% of the SIM principal respondents rated that the SIM programme "useful" or "very useful."

Descriptive Analysis – Measure of Central Tendency

Table 274: Results of the SIM principals' SIM usefulness rating frequency distribution

```
. tabulate q10

        q10 |      Freq.     Percent        Cum.
------------+-----------------------------------
          2 |         11        8.94        8.94
          3 |         64       52.03       60.98
          4 |         48       39.02      100.00
------------+-----------------------------------
      Total |        123      100.00
```

From the frequency ***Table 274*** above, it shows that mode choice is 3, which is "useful." The total SIM principal respondents of 91.0% chose "useful" or "very useful" in their perception on usefulness of SIM.

Table 275: SIM principals' SIM usefulness rating frequency distribution, by school type

```
. tabulate school q10

           |               q10
    School |         2          3          4 |     Total
-----------+---------------------------------+----------
       ECR |         1          3          2 |         6
       HSS |         6         13          4 |        23
       LSS |         0          8          1 |         9
       MSS |         0          8          6 |        14
        PS |         4         32         35 |        71
-----------+---------------------------------+----------
     Total |        11         64         48 |       123
```

Looking at principals' perception on usefulness of SIM by school type, it shows that in majority school types, the mode is 3, which is "useful" and in the case of primary schools, the mode is 4, which is "very useful."

Table 276: Result of the SIM principals' SIM usefulness rating median calculation

```
. tabstat q10, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q10 |       123         3         2         4
------------------------------------------------------
```

The calculated sample median = 3, which is "useful." This means at least 50% of the SIM principal respondents believe that SIM was "useful" or "very useful" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 277: Result of the SIM Principals' measure of consensus on SIM usefulness

```
. cns q10 , min(1) max(4)

Consensus Measure for q10
Cns(X) = .69794634
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the SIM usefulness opinion of SIM principals, it is 0.6979.

Inferential Analysis – Statistical Significance Testing

Table 278: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q10 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      112        7208        3813
    Negative |       11         418        3813
        Zero |        0           0           0
-------------+---------------------------------
         All |      123        7626        7626

Unadjusted variance   156968.50
Adjustment for ties   -11090.50
Adjustment for zeros       0.00
                     ----------
Adjusted variance     145878.00

H0: q10 = 2.5
         z =  8.889
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 91.0% of SIM principals surveyed think that SIM programme was useful. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM principal population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly useful" and 3 = "useful."

Ho: Our null hypothesis is that the SIM principal population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM principal population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.889, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 8.889 and our sample size for SIM Principals is 123. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 8.889/SQRT(123) = 0.80. This, according to Bartz (1999), is very strong effect size.

Evidence on SIM Principals' Perception on Usefulness of SIM

There is statistically significant evidence (p = 0.0000) that 91.0% of SIM principals believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 8.889, p = 0.0000, with a very strong effect size (*r* = 0.80).

Analyzing Principals’ Perception on Overall Presentation of SIM

The Ministry of Education was interested to know how attractive was overall presentation of SIM booklets in the perception of principals. To investigate this, ***Figure 61***, which is visualization of survey data, shows the results of overall presentation of SIM booklets in the perception of principals.

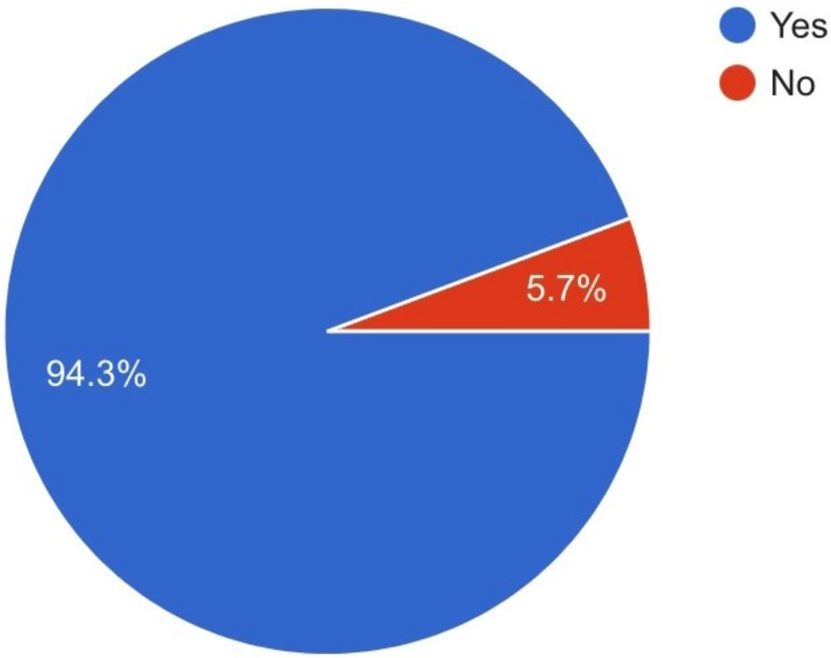


Figure 61: *Results of "Is overall presentation of SIM attractive?"*

As can be seen in ***Figure 61*** the 94.3% of the SIM principal respondents rated that overall presentation of SIM booklets is attractive.

Inferential Analysis – Statistical Significance Testing

Table 279: Results of Binomial Test on Principals' Perception on Overall Presentation of SIM

```
. bitest q9a = 0.88

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q9a |      123          116       108.24     0.88000      0.94309

  Pr(k >= 116)           = 0.015409  (one-sided test)
  Pr(k <= 116)           = 0.993831  (one-sided test)
  Pr(k <= 99 or k >= 116) = 0.026397  (two-sided test
```

A binomial test indicated that the percentage of SIM principals who believe overall presentation of SIM is attractive (N_{Yes} = 116, 94.3%), was statistically significantly greater than the population hypothesized value of 88%, p = 0.015409.

Evidence on SIM Principals' Perception on Overall Presentation of SIM

There is statistically significant evidence (p = 0.015409) that at least 88% of SIM principals believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM principals who believe overall presentation of SIM is attractive (N_{Yes} = 116, 94.3%), was statistically significantly greater than the population hypothesized value of 88%, p = 0.015409.

Analyzing Principals' Perception on Schools' Support Extended to SIM Students

The Ministry of Education was interested to know about support extended to SIM students in the perception of principals. To investigate this, ***Figure 62***, which is visualization of survey data, shows the results of support extended to SIM students in the perception of principals.

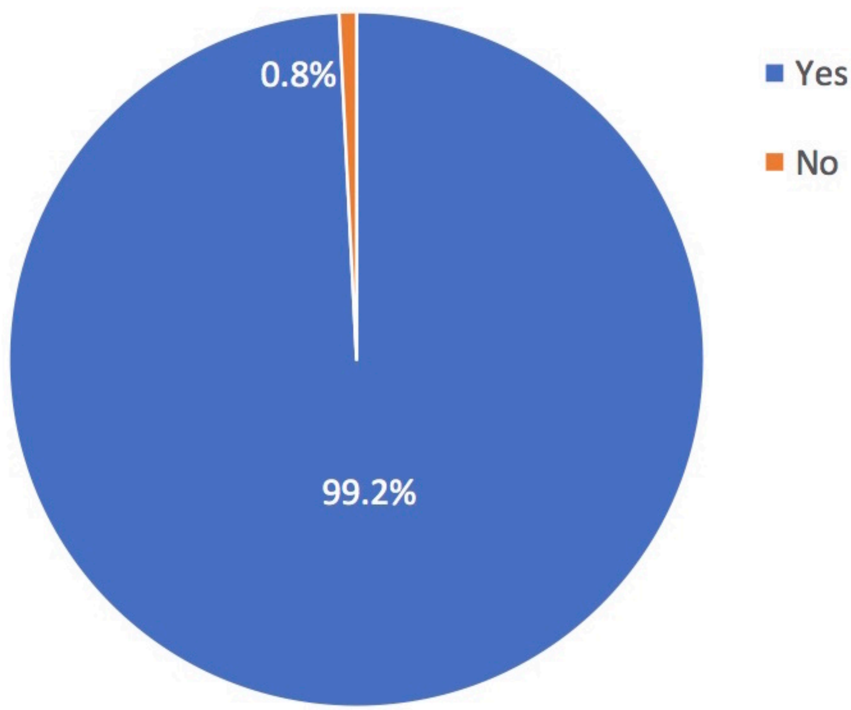


Figure 62*: Results of "Did the school extend support to the students?"*

As can be seen in ***Figure 62*** the 99.2% of the SIM principal respondents rated that their schools extended support to the SIM students.

Inferential Analysis – Statistical Significance Testing

Table 280: Results of Binomial Test on Support Extended to SIM Students

```
. bitest q8a = 0.95

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q8a |      123         122       116.85     0.95000      0.99187

  Pr(k >= 122)            = 0.013600  (one-sided test)
  Pr(k <= 122)            = 0.998180  (one-sided test)
  Pr(k <= 110 or k >= 122) = 0.022393  (two-sided test)
```

A binomial test indicated that the percentage of SIM principals who believe their schools extended support to SIM students (N_{Yes} = 122, 99.2%), was statistically significantly greater than the population hypothesized value of 95%, p = 0.013600.

Evidence on SIM Principals' Perception on Support Extended to SIM Students

There is statistically significant evidence (p = 0.013600) that at least 95% of SIM schools extended support to SIM students. A binomial test indicated that the percentage of SIM principals who believe their schools extended support to SIM students (N_{Yes} = 122, 99.2%), was statistically significantly greater than the population hypothesized value of 95%, p = 0.013600.

Analyzing Principals' Perception on Help Sought by SIM Students and Parents

The Ministry of Education was interested to know about help sought by SIM students and parents in the perception of principals. To investigate this, ***Figure 63***, which is visualization of survey data, shows the results of help sought by SIM students and parents in the perception of principals.

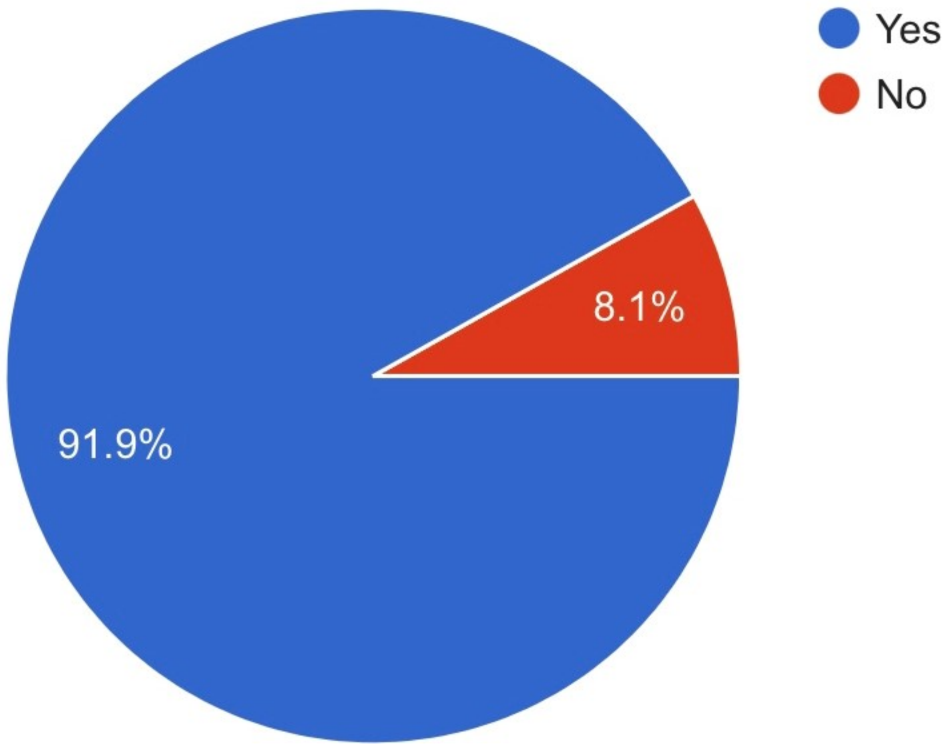


Figure 63: *Results of "Did your students/parents seek any help regarding SIM?"*

As can be seen in ***Figure 63*** the 91.9% of the SIM principal respondents said that their students or students' parents sought help regarding SIM.

Inferential Analysis – Statistical Significance Testing

Table 281: Results of Binomial Test on Help Sought by SIM Students and Parents

```
. bitest q7a = 0.85

Binomial probability test

Variable |         N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q7a |       123          113       104.55     0.85000      0.91870

  Pr(k >= 113)            = 0.016869  (one-sided test)
  Pr(k <= 113)            = 0.992280  (one-sided test)
  Pr(k <= 95 or k >= 113) = 0.031314  (two-sided test)
```

A binomial test indicated that the percentage of SIM principals who believe their students or students' parents sought help regarding SIM (N_{Yes} = 113, 91.9%), was statistically significantly greater than the population hypothesized value of 85%, $p = 0.016869$.

Evidence on SIM Principals' Perception on Help Sought by SIM Students and Parents

There is statistically significant evidence (p = 0.016869) that at least 85% of SIM students and parents sought help regarding SIM in the perception of principals. A binomial test indicated that the percentage of SIM principals who believe their students or students' parents sought help regarding SIM (N_{Yes} = 113, 91.9%), was statistically significantly greater than the population hypothesized value of 85%, $p = 0.016869$.

Analyzing Principals' Perception on Whether DEOs Delivered SIMs

The Ministry of Education was interested to know whether the Dzongkhag Education Office delivered the SIMs, in the perception of principals. To investigate this, ***Figure 64**,* which is visualization of survey data, shows the results of whether DEOs delivered SIMs or not, in the perception of principals.

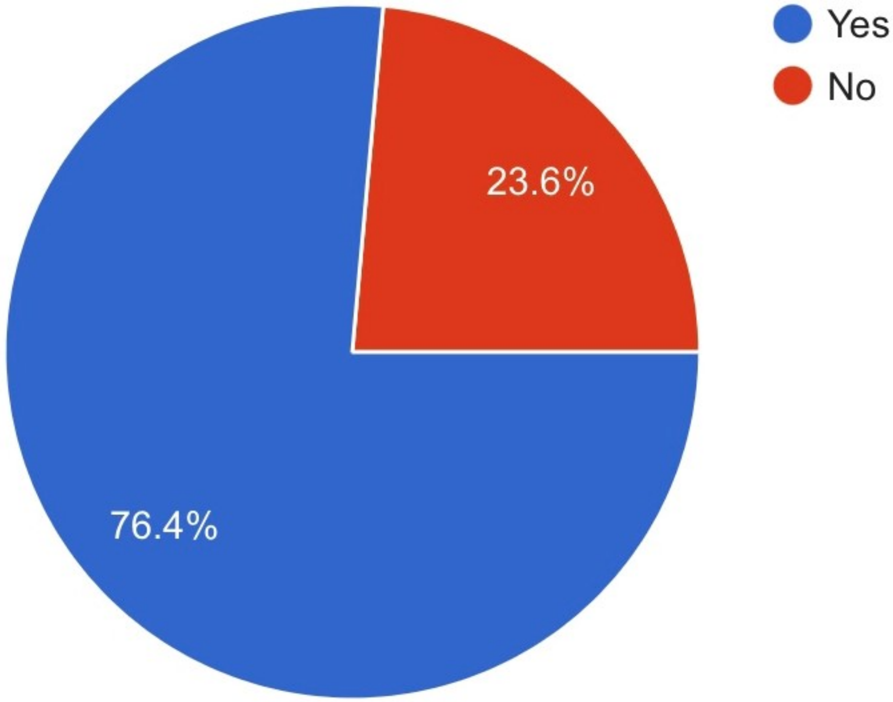


***Figure 64**: Results of "Did the Dzongkhag Education Office deliver the SIMs?"*

As can be seen in ***Figure 64*** the 76.4% of the SIM principal respondents said that the Dzongkhag Education Office delivered the SIMs.

Inferential Analysis – Statistical Significance Testing

Table 282: Results of Binomial Test on Whether DEOs Delivered the SIMs

```
. bitest q6a = 0.675

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q6a |        123           94       83.025     0.67500      0.76423

  Pr(k >= 94)            = 0.019772  (one-sided test)
  Pr(k <= 94)            = 0.988184  (one-sided test)
  Pr(k <= 71 or k >= 94) = 0.034303  (two-sided test)
```

A binomial test indicated that the percentage of SIM principals who believe the Dzongkhag Education Office delivered the SIMs (N_{Yes} = 94, 76.4%), was statistically significantly greater than the population hypothesized value of 67.5%, p = 0.019772.

Evidence on SIM Principals' Perception on Whether DEOs Delivered the SIMs

There is statistically significant evidence (p = 0.019772) that at least 67.5% of SIM principals believe the Dzongkhag Education Office delivered the SIMs. A binomial test indicated that the percentage of SIM principals who believe the Dzongkhag Education Office delivered the SIMs (N_{Yes} = 94, 76.4%), was statistically significantly greater than the population hypothesized value of 67.5%, p = 0.019772.

Analyzing Principals' Perception on Whether SIM Reached the Identified Students

The Ministry of Education was interested to know whether the SIM reached the identified students. To investigate this, ***Figure 65***, which is visualization of survey data, shows the results of whether SIM reached the identified students or not, in the perception of principals.

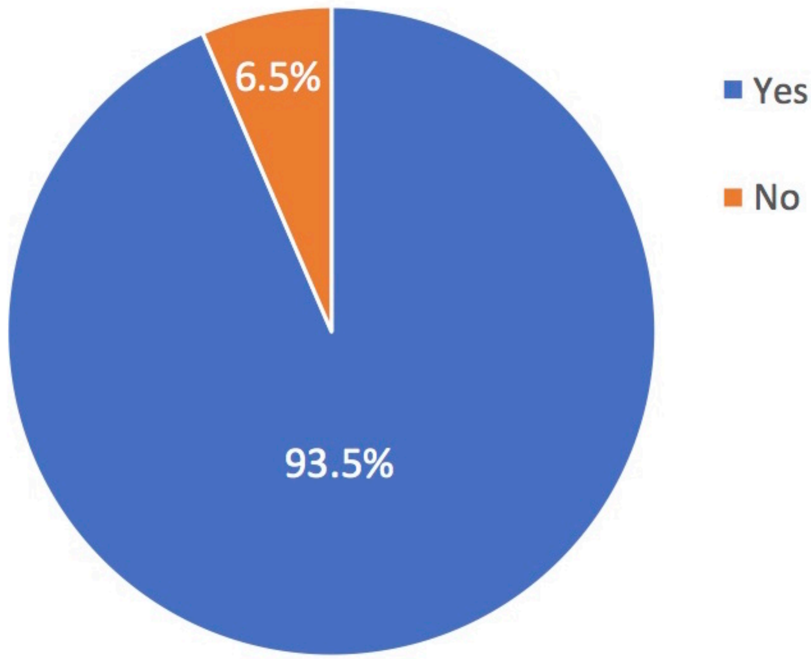


Figure 65: *Results of "Has the SIM reached the identified students?"*

As can be seen in ***Figure 65*** the 93.5% of the SIM principal respondents said that the SIM has reached the identified students.

Inferential Analysis – Statistical Significance Testing

Table 283: Results of Binomial Test on Whether SIM Reached the Identified Students

```
. bitest q4a = 0.875

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q4a |        123          115      107.625     0.87500      0.93496

  Pr(k >= 115)            = 0.023463  (one-sided test)
  Pr(k <= 115)            = 0.989689  (one-sided test)
  Pr(k <= 99 or k >= 115) = 0.040927  (two-sided test)
```

A binomial test indicated that the percentage of SIM principals who believe the SIM has reached the identified students (N_{Yes} = 115, 93.5%), was statistically significantly greater than the population hypothesized value of 87.5%, p = 0.023463.

Evidence on SIM Principals' Perception on Whether SIM Reached the Identified Students

There is statistically significant evidence (p = 0.023463) that at least 87.5% of SIM principals believe SIM has reached the identified students. A binomial test indicated that the percentage of SIM principals who believe the SIM has reached the identified students (N_{Yes} = 115, 93.5%), was statistically significantly greater than the population hypothesized value of 87.5%, p = 0.023463.

Analyzing Principals' Perception on Whether SIM Reached Other Needy Students

The Ministry of Education was interested to know whether the SIM reached other needy students beyond the identified students. To investigate this, ***Figure 66***, which is visualization of survey data, shows the results of whether SIM reached other needy students beyond the identified students or not, in the perception of principals.

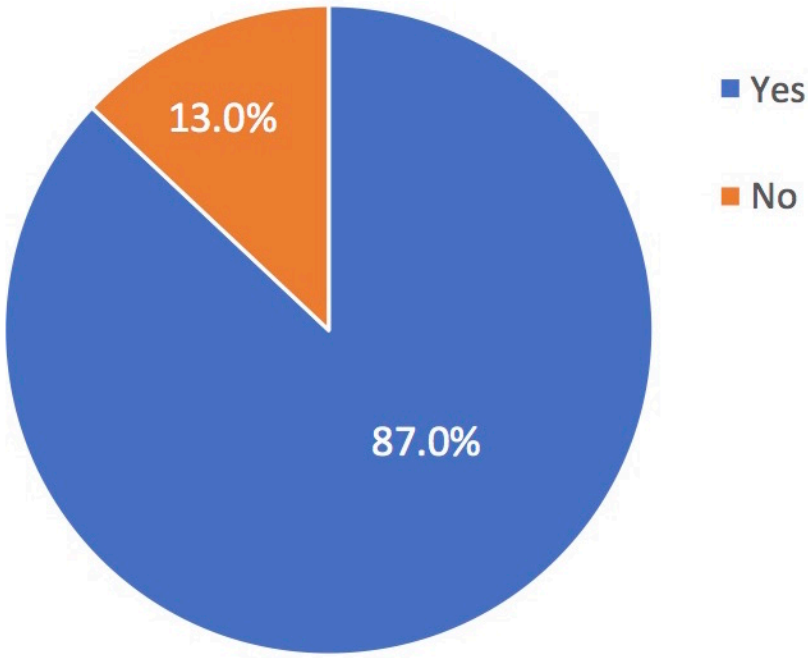


Figure 66: *Results of "Has the SIM reached other needy students beyond the identified students?"*

As can be seen in ***Figure 66*** the 87.0% of the SIM principal respondents said that the SIM has reached other needy students.

Inferential Analysis – Statistical Significance Testing

Table 284: Results of Binomial Test on Whether SIM Reached Other Needy Students

```
. bitest q5a = 0.795

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q5a |       123          107       97.785     0.79500      0.86992

  Pr(k >= 107)            = 0.021581  (one-sided test)
  Pr(k <= 107)            = 0.988477  (one-sided test)
  Pr(k <= 88 or k >= 107) = 0.043620  (two-sided test)
```

A binomial test indicated that the percentage of SIM principals who believe the SIM has reached other needy students (N_{Yes} = 107, 87.0%), was statistically significantly greater than the population hypothesized value of 79.5%, p = 0.021581.

Evidence on SIM Principals' Perception on Whether SIM Reached Other Needy Students

There is statistically significant evidence (p = 0.021581) that at least 79.5% of SIM principals believe SIM has reached other needy students. A binomial test indicated that the percentage of SIM principals who believe the SIM has reached other needy students (N_{Yes} = 107, 87.0%), was statistically significantly greater than the population hypothesized value of 79.5%, p = 0.021581.

PART IV: SIM District Education Officers

Demographic Characteristics of SIM DEO Respondents

The age characteristics of the SIM DEO respondents are summarized in ***Table 285***. The age of the SIM DEO respondents ranged from 41 to 54 years (M = 48.24, SD = 4.09).

Table 285: Results of age characteristics of SIM DEO respondents

Variable	Obs	Mean	Std. dev.	Min	Max
age	29	48.24138	4.085297	41	54

Similarly, among the 29 SIM chief DEO and deputy DEO respondents, 26 (89.7%) were males and 3 (10.3%) were females as shown in ***Figure 67***. The low representation of female DEOs is the reflection of reality in the population through random sampling and not lack of data.

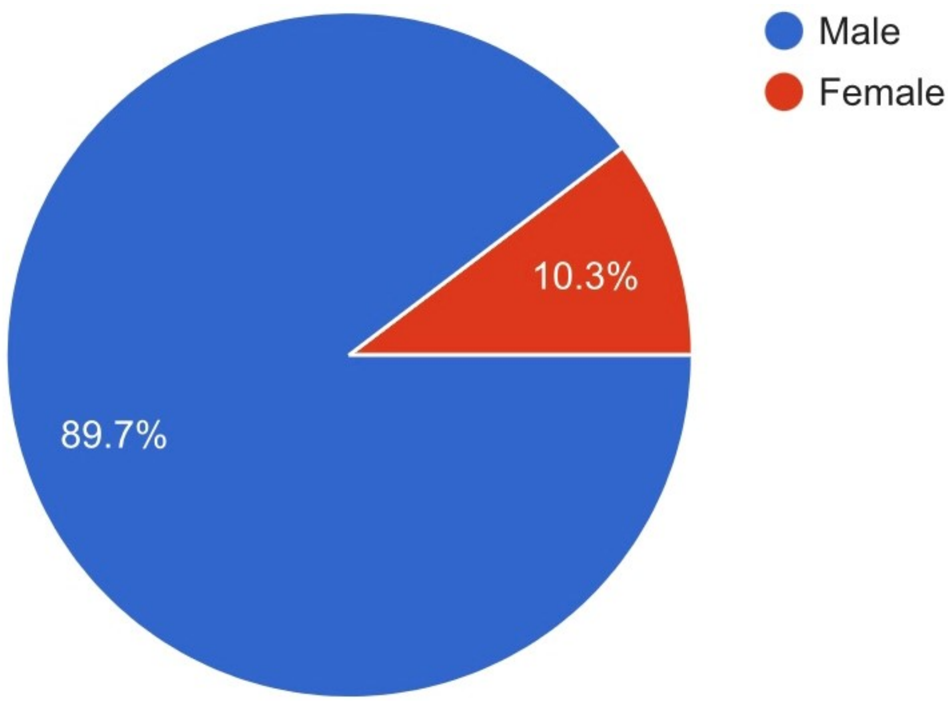


Figure 67: *Gender of SIM DEO respondents*

Effectiveness of SIM Programme

Analyzing DEOs' Satisfaction Level of SIM

The Ministry of Education was interested to know satisfaction level of SIM programme, including DEOs' satisfaction level, during COVID-19 pandemic. To investigate this, ***Figure 68***, which is visualization of survey data, shows the results of satisfaction level of DEOs from the SIM survey.

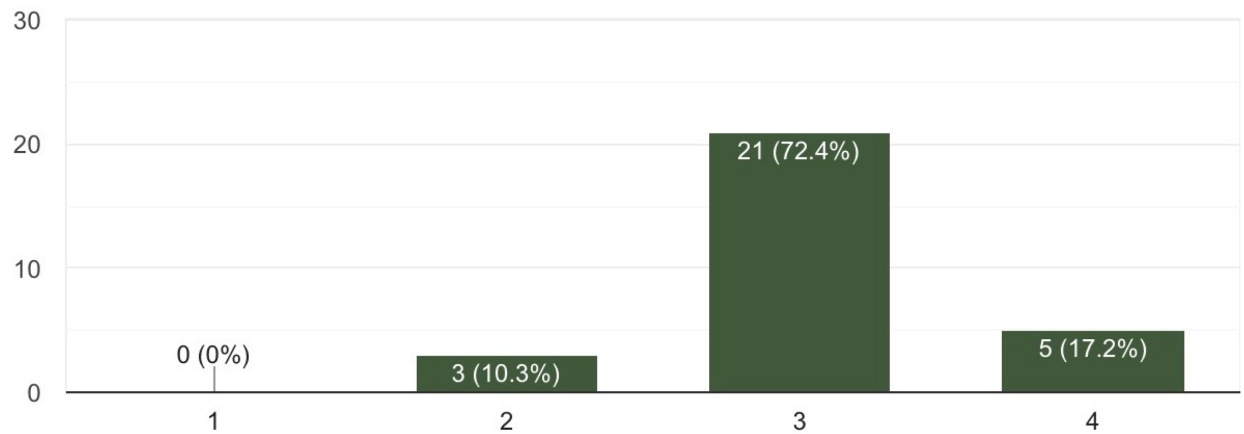


Figure 68*: Results of "Rate how satisfied are you with the current SIM" where 1 = Extremely dissatisfied, 2 = Dissatisfied, 3 = Satisfied, and 4 = Extremely satisfied*

As can be seen in ***Figure 68*** the 89.6% of the SIM DEO respondents rated the SIM programme "satisfied" or "extremely satisfied."

Descriptive Analysis – Measure of Central Tendency

Table 286: Results of the SIM DEOs' satisfaction level rating frequency distribution

```
. tabulate q9

         q9 |      Freq.     Percent        Cum.
------------+-----------------------------------
          2 |          3       10.34       10.34
          3 |         21       72.41       82.76
          4 |          5       17.24      100.00
------------+-----------------------------------
      Total |         29      100.00
```

From the frequency ***Table 286*** above, it shows that mode choice is 3, which is "satisfied." The total SIM DEO respondents of 89.6% chose "satisfied" or "extremely satisfied."

Table 287: Result of the SIM DEOs' satisfaction level rating median calculation

```
. tabstat q9, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q9 |        29         3         2         4
------------------------------------------------------
```

The calculated sample median = 3, which is "satisfied." This means at least 50% of the SIM DEO respondents are in the "satisfied" or "extremely satisfied" category looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 288: Result of the SIM DEOs' measure of consensus on satisfaction level

```
. cns q9 , min(1) max(4)

Consensus Measure for q9
Cns(X) = .81752987
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the satisfaction level of SIM DEOs, it is 0.8175.

Inferential Analysis – Statistical Significance Testing

Table 289: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q9 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       26       397.5       217.5
    Negative |        3        37.5       217.5
        Zero |        0           0           0
-------------+---------------------------------
         All |       29         435         435

Unadjusted variance     2138.75
Adjustment for ties     -290.00
Adjustment for zeros       0.00
                     ----------
Adjusted variance       1848.75

H0: q9 = 2.5
         z =  4.186
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 89.6% of SIM DEOs surveyed think that SIM programme was satisfactory. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM DEO population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "dissatisfied" and 3 = "satisfied."

Ho: Our null hypothesis is that the SIM DEO population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM DEO population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.186, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 4.186 and our sample size for SIM DEOs is 29. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 4.186/SQRT(29) = 0.78. This, according to Bartz (1999), is strong effect size.

Evidence on SIM DEOs' Satisfaction Level

There is statistically significant evidence (p = 0.0000) that 89.6% of SIM DEOs are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.186, p = 0.0000, with a strong effect size (*r* = 0.78).

Analyzing DEOs' Perception on Implementation of SIM

The Ministry of Education was interested to know how effectively implementation of SIM programme was carried out in the perception of DEOs. To investigate this, ***Figure 69***, which is visualization of survey data, shows the results of implementation effectiveness perception of DEOs from the SIM survey.

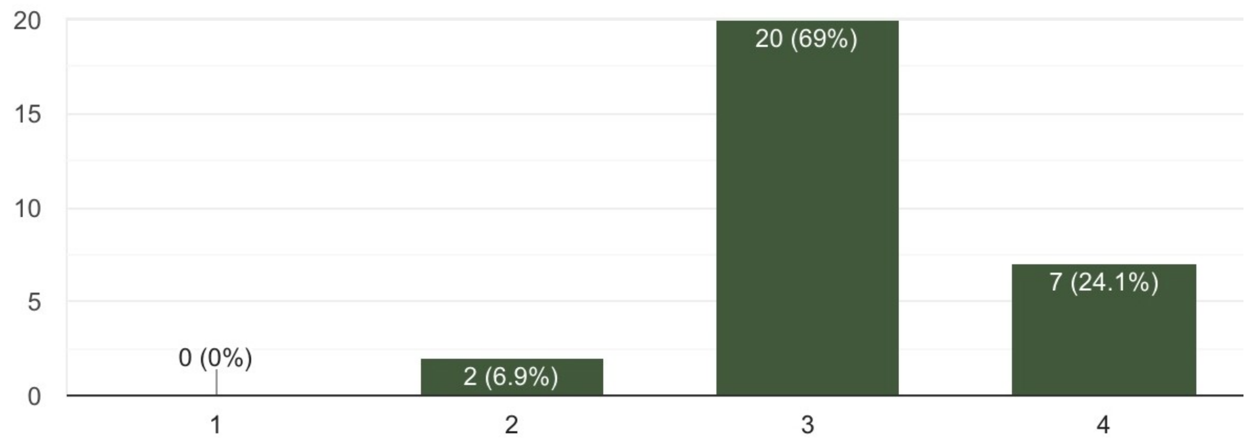


Figure 69*: Results of "Rate how effectively has the SIM been implemented" where 1 = Not effective, 2 = Slightly effective, 3 = Effective, and 4 = Very effective*

As can be seen in ***Figure 69*** the 93.1% of the SIM DEO respondents rated that the SIM programme implementation was "effective" or "very effective."

Descriptive Analysis – Measure of Central Tendency

Table 290: Results of the SIM DEOs' SIM implementation rating frequency distribution

```
. tabulate q8

         q8 |      Freq.     Percent        Cum.
------------+-----------------------------------
          2 |          2        6.90        6.90
          3 |         20       68.97       75.86
          4 |          7       24.14      100.00
------------+-----------------------------------
      Total |         29      100.00
```

From the frequency ***Table 290*** above, it shows that mode choice is 3, which is "effective." The total SIM DEO respondents of 93.1% chose "effective" or "very effective" in their perception on implementation effectiveness of the SIM.

Table 291: Result of the SIM DEOs' SIM implementation rating median calculation

```
. tabstat q8, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q8 |        29         3         2         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM DEO respondents believe that SIM implementation was "effective" or "very effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 292: Result of the SIM DEOs' measure of consensus on SIM implementation

```
. cns q8 , min(1) max(4)

Consensus Measure for q8
Cns(X) = .7793958
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the implementation effectiveness opinion of SIM DEOs, it is 0.7794.

Inferential Analysis – Statistical Significance Testing

Table 293: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       27         412       217.5
    Negative |        2          23       217.5
        Zero |        0           0           0
-------------+---------------------------------
         All |       29         435         435

Unadjusted variance     2138.75
Adjustment for ties     -228.38
Adjustment for zeros       0.00
                     ----------
Adjusted variance       1910.38

H0: q8 = 2.5
         z =  4.450
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 93.1% of SIM DEOs surveyed think that SIM programme was effectively implemented. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM DEO population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly effective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM DEO population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM DEO population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.450, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 4.450 and our sample size for SIM DEOs is 29. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 4.45/SQRT(29) = 0.83. This, according to Bartz (1999), is very strong effect size.

Evidence on SIM DEOs' Perception on Implementation Effectiveness of SIM

There is statistically significant evidence (p = 0.0000) that 93.1% of SIM DEOs believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.450, p = 0.0000, with a very strong effect size (r = 0.83).

Analyzing DEOs' Perception on Usefulness of SIM

The Ministry of Education was interested to know how useful was SIM programme in the perception of DEOs. To investigate this, ***Figure 70***, which is visualization of survey data, shows the results of SIM usefulness perception of DEOs from the SIM survey.

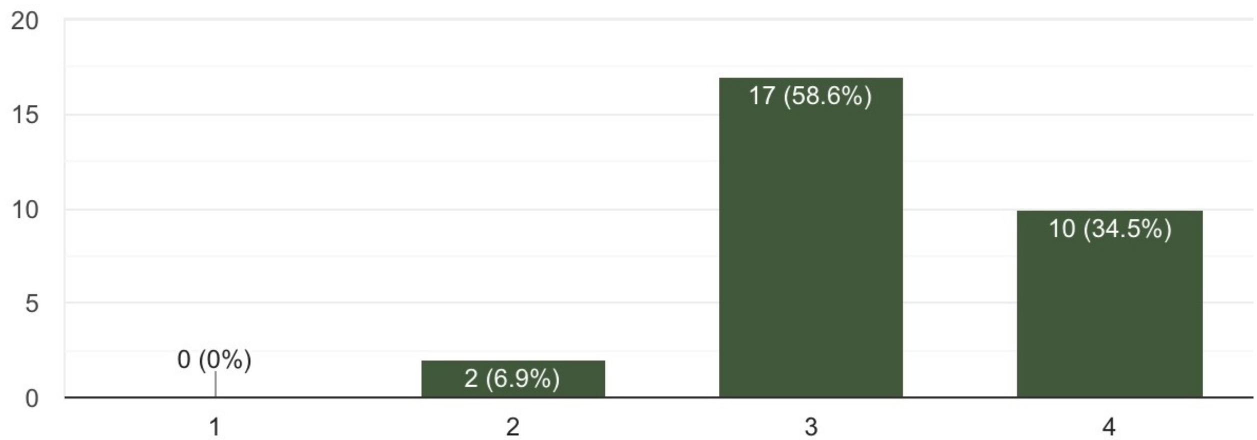


Figure 70*: Results of "Rate how useful was SIM" where 1 = Not useful, 2 = Slightly useful, 3 = Useful, and 4 = Very useful*

As can be seen in ***Figure 70*** the 93.1% of the SIM DEO respondents rated that the SIM programme "useful" or "very useful."

Descriptive Analysis – Measure of Central Tendency

Table 294: Results of the SIM DEOs' SIM usefulness rating frequency distribution

```
. tabulate q7

         q7 |      Freq.     Percent        Cum.
------------+-----------------------------------
          2 |          2        6.90        6.90
          3 |         17       58.62       65.52
          4 |         10       34.48      100.00
------------+-----------------------------------
      Total |         29      100.00
```

From the frequency ***Table 294*** above, it shows that mode choice is 3, which is "useful." The total SIM DEO respondents of 93.1% chose "useful" or "very useful" in their perception on usefulness of SIM.

Table 295: Result of the SIM DEOs' SIM usefulness rating median calculation

```
. tabstat q7, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q7 |        29         3         2         4
------------------------------------------------------
```

The calculated sample median = 3, which is "useful." This means at least 50% of the SIM DEO respondents believe that SIM was "useful" or "very useful" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 296: Result of the SIM DEOs' measure of consensus on SIM usefulness

```
. cns q7 , min(1) max(4)

Consensus Measure for q7
Cns(X) = .72588152
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the SIM usefulness opinion of SIM DEOs, it is 0.7259.

Inferential Analysis – Statistical Significance Testing

Table 297: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q7 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       27         415       217.5
    Negative |        2          20       217.5
        Zero |        0           0           0
-------------+---------------------------------
         All |       29         435         435

Unadjusted variance     2138.75
Adjustment for ties     -163.13
Adjustment for zeros       0.00
                     ----------
Adjusted variance       1975.63

H0: q7 = 2.5
         z =  4.443
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 93.1% of SIM DEOs surveyed think that SIM programme was useful. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM DEO population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly useful" and 3 = "useful."

Ho: Our null hypothesis is that the SIM DEO population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM DEO population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 4.443, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 4.443 and our sample size for SIM DEOs is 29. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 4.443/SQRT(29) = 0.83. This, according to Bartz (1999), is very strong effect size.

Evidence on SIM DEOs' Perception on Usefulness of SIM

There is statistically significant evidence ($p = 0.0000$) that 93.1% of SIM DEOs believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 4.443$, $p = 0.0000$, with a very strong effect size ($r = 0.83$).

Analyzing DEOs' Perception on Overall Presentation of SIM

The Ministry of Education was interested to know how attractive was overall presentation of SIM booklets in the perception of DEOs. To investigate this, ***Figure 71***, which is visualization of survey data, shows the results of overall presentation of SIM booklets in the perception of DEOs.

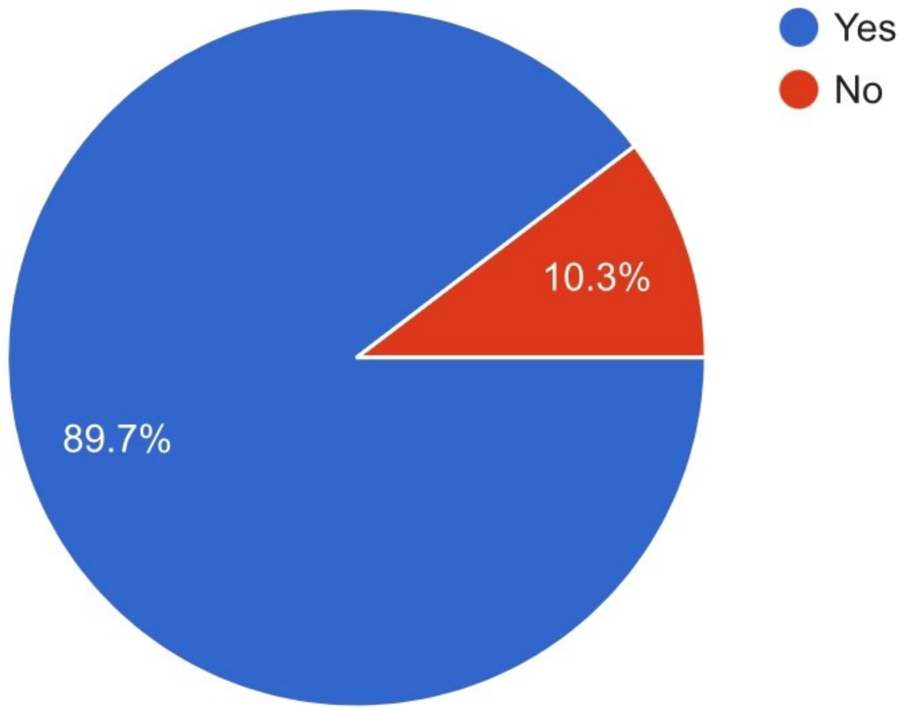


Figure 71*: Results of "Is overall presentation of SIM attractive?"*

As can be seen in ***Figure 71*** the 89.7% of the SIM DEO respondents rated that overall presentation of SIM booklets is attractive.

Inferential Analysis – Statistical Significance Testing

Table 298: Results of Binomial Test on DEOs' Perception on Overall Presentation of SIM

```
. bitest q6a = 0.74

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q6a |         29           26        21.46     0.74000      0.89655

  Pr(k >= 26)            = 0.035460  (one-sided test)
  Pr(k <= 26)            = 0.990109  (one-sided test)
  Pr(k <= 16 or k >= 26) = 0.057258  (two-sided test)
```

A binomial test indicated that the percentage of SIM DEOs who believe overall presentation of SIM is attractive (N_{Yes} = 26, 89.7%), was statistically significantly greater than the population hypothesized value of 74%, p = 0.035460.

Evidence on SIM DEOs' Perception on Overall Presentation of SIM

There is statistically significant evidence (p = 0.0345460) that at least 74% of SIM DEOs believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM DEOs who believe overall presentation of SIM is attractive (N_{Yes} = 26, 89.7%), was statistically significantly greater than the population hypothesized value of 74%, p = 0.035460.

Analyzing DEOs' Perception on Whether They Delivered SIMs

The Ministry of Education was interested to know whether the Dzongkhag Education Office delivered the SIMs, in the perception of DEOs. To investigate this, ***Figure 72***, which is visualization of survey data, shows the results of whether DEOs delivered SIMs or not, in the perception of DEOs.

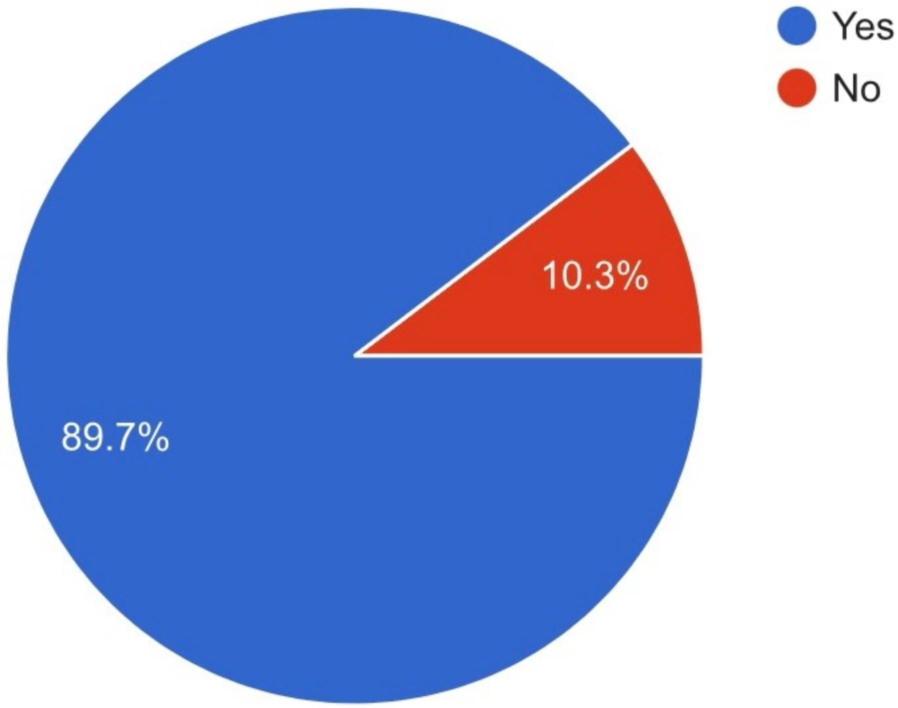


***Figure 72**: Results of "Did the Dzongkhag Education Office deliver the SIMs?"*

As can be seen in ***Figure 72*** the 89.7% of the SIM DEO respondents said that the Dzongkhag Education Office delivered the SIMs.

Inferential Analysis – Statistical Significance Testing

Table 299: Results of Binomial Test on Whether DEOs Delivered the SIMs

```
. bitest q5a = 0.74

Binomial probability test

Variable |         N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q5a |        29           26        21.46     0.74000      0.89655

  Pr(k >= 26)           = 0.035460  (one-sided test)
  Pr(k <= 26)           = 0.990109  (one-sided test)
  Pr(k <= 16 or k >= 26) = 0.057258  (two-sided test)
```

A binomial test indicated that the percentage of SIM DEOs who believe the Dzongkhag Education Office delivered the SIMs (N_{Yes} = 26, 89.7%), was statistically significantly greater than the population hypothesized value of 74%, p = 0.035460.

Evidence on SIM DEOs' Perception on Whether DEOs Delivered the SIMs

There is statistically significant evidence (p = 0.0035460) that at least 74% of SIM DEOs believe the Dzongkhag Education Office delivered the SIMs. A binomial test indicated that the percentage of SIM DEOs who believe the Dzongkhag Education Office delivered the SIMs (N_{Yes} = 26, 89.7%), was statistically significantly greater than the population hypothesized value of 74%, p = 0.035460.

Analyzing DEOs' Perception on Whether SIM Reached the Identified Students

The Ministry of Education was interested to know whether the SIM reached the identified students. To investigate this, ***Figure 73,*** which is visualization of survey data, shows the results of whether SIM reached the identified students or not, in the perception of DEOs.

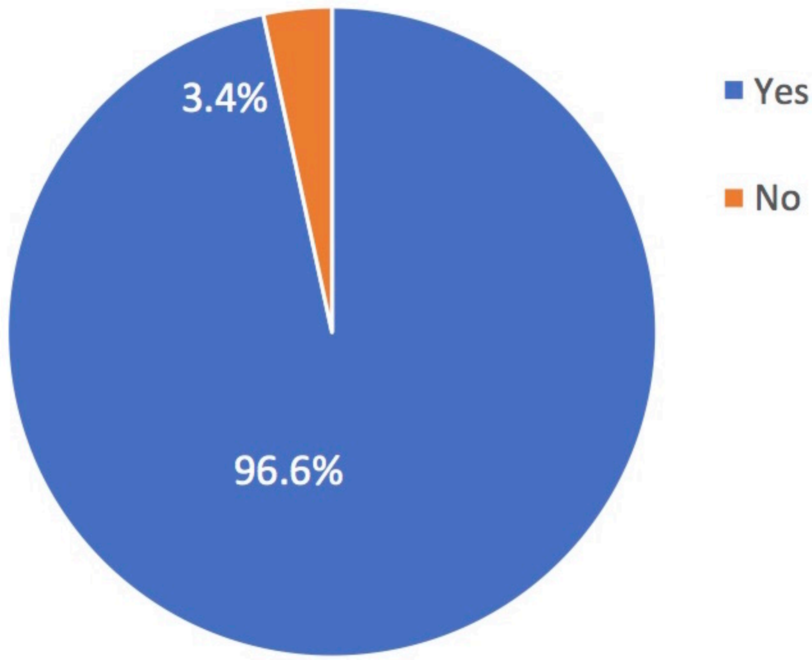


Figure 73*: Results of "Has the SIM reached the identified students?"*

As can be seen in ***Figure 73*** the 96.6% of the SIM DEO respondents said that the SIM has reached the identified students.

Inferential Analysis – Statistical Significance Testing

Table 300: Results of Binomial Test on Whether SIM Reached the Identified Students

```
. bitest q3a = 0.84

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+-------------------------------------------------------------
     q3a |         29           28        24.36     0.84000      0.96552

  Pr(k >= 28)            = 0.041553  (one-sided test)
  Pr(k <= 28)            = 0.993631  (one-sided test)
  Pr(k <= 20 or k >= 28) = 0.074257  (two-sided test)
```

A binomial test indicated that the percentage of SIM DEOs who believe the SIM has reached the identified students (N_{Yes} = 28, 96.6%), was statistically significantly greater than the population hypothesized value of 84%, $p = 0.041553$.

Evidence on SIM DEOs' Perception on Whether SIM Reached the Identified Students

There is statistically significant evidence (p = 0.041553) that at least 84% of SIM DEOs believe SIM has reached the identified students. A binomial test indicated that the percentage of SIM DEOs who believe the SIM has reached the identified students (N_{Yes} = 28, 96.6%), was statistically significantly greater than the population hypothesized value of 84%, $p = 0.041553$.

Analyzing DEOs' Perception on Whether SIM Reached Other Needy Students

The Ministry of Education was interested to know whether the SIM reached other needy students beyond the identified students. To investigate this, ***Figure 74***, which is visualization of survey data, shows the results of whether SIM reached other needy students beyond the identified students or not, in the perception of DEOs.

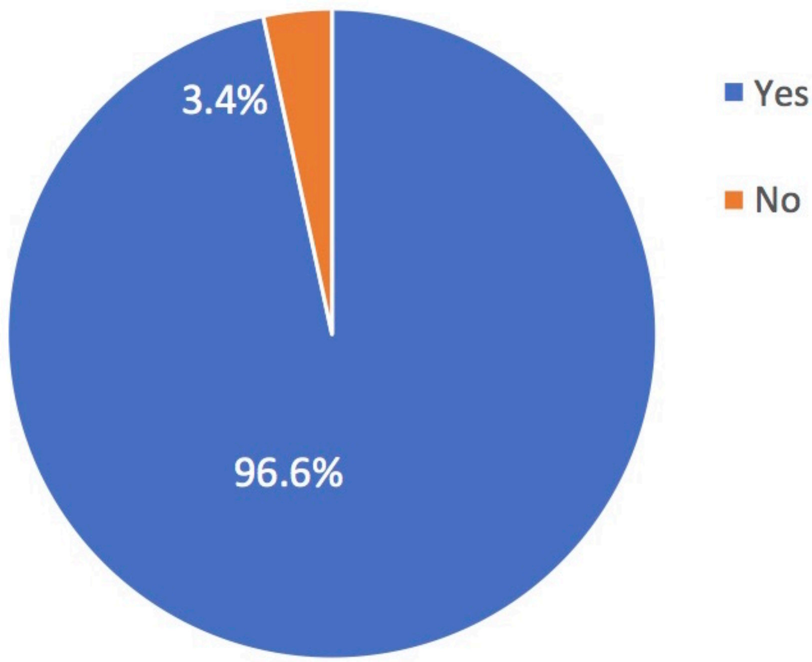


Figure 74: *Results of "Has the SIM reached other needy students beyond the identified students?"*

As can be seen in ***Figure 74*** the 96.6% of the SIM DEO respondents said that the SIM has reached other needy students.

Inferential Analysis – Statistical Significance Testing

Table 301: Results of Binomial Test on Whether SIM Reached Other Needy Students

```
. bitest q4a = 0.84

Binomial probability test

Variable |         N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q4a |        29          28        24.36     0.84000      0.96552

  Pr(k >= 28)            = 0.041553  (one-sided test)
  Pr(k <= 28)            = 0.993631  (one-sided test)
  Pr(k <= 20 or k >= 28) = 0.074257  (two-sided test)
```

A binomial test indicated that the percentage of SIM DEOs who believe the SIM has reached other needy students (N_{Yes} = 28, 96.6%), was statistically significantly greater than the population hypothesized value of 84%, p = 0.041553.

Evidence on SIM DEOs' Perception on Whether SIM Reached Other Needy Students

There is statistically significant evidence (p = 0.041553) that at least 84% of SIM DEOs believe SIM has reached other needy students. A binomial test indicated that the percentage of SIM DEOs who believe the SIM has reached other needy students (N_{Yes} = 28, 96.6%), was statistically significantly greater than the population hypothesized value of 84%, p = 0.041553.

PART V: SIM LOCAL GOVERNMENT LEADERS

Demographic Characteristics of SIM LG Respondents

The age characteristics of the SIM LG respondents are summarized in ***Table 302***. The age of the SIM LG respondents ranged from 27 to 58 years (M = 37.67, SD = 6.82).

Table 302: Results of age characteristics of SIM LG respondents

Variable	Obs	Mean	Std. dev.	Min	Max
age	76	37.67105	6.822293	27	58

Similarly, among the 76 SIM LG respondents, 65 (85.5%) were males and 11 (14.5%) were females as shown in ***Figure 75***. The low representation of female LG leaders is the reflection of reality in the population through random sampling and not lack of data.

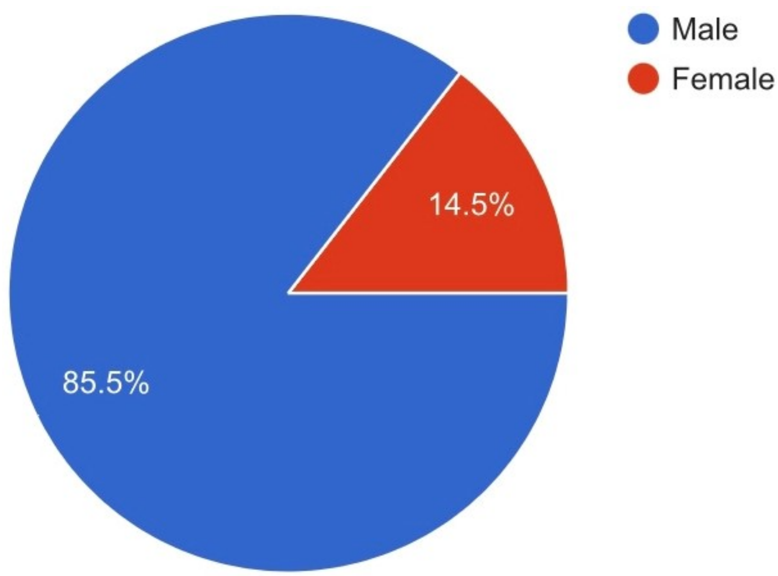


Figure 75: *Gender of SIM LG respondents*

Effectiveness of SIM Programme

Analyzing LG leaders' Satisfaction Level of SIM

The Ministry of Education was interested to know satisfaction level of SIM programme, including LG leaders' satisfaction level, during COVID-19 pandemic. To investigate this, ***Figure 76***, which is visualization of survey data, shows the results of satisfaction level of LG leaders from the SIM survey.

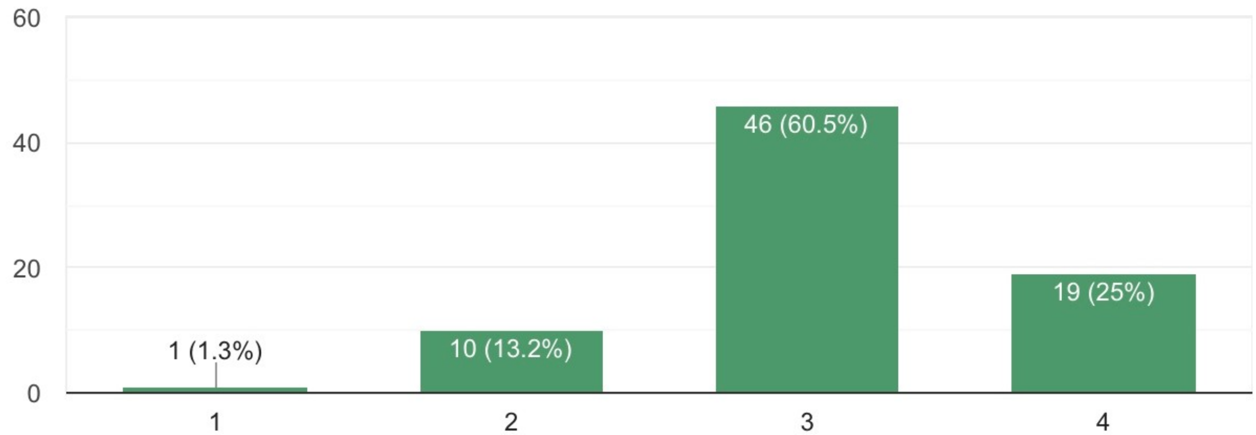


***Figure 76**: Results of "Rate how satisfied are you with the current SIM" where 1 = Extremely dissatisfied, 2 = Dissatisfied, 3 = Satisfied, and 4 = Extremely satisfied*

As can be seen in ***Figure 76*** the 85.5% of the SIM LG leader respondents rated the SIM programme "satisfied" or "extremely satisfied."

Descriptive Analysis – Measure of Central Tendency

Table 303: Results of the SIM LG leaders' satisfaction level rating frequency distribution

```
. tabulate q9

         q9 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          1        1.32        1.32
          2 |         10       13.16       14.47
          3 |         46       60.53       75.00
          4 |         19       25.00      100.00
------------+-----------------------------------
      Total |         76      100.00
```

From the frequency ***Table 303*** above, it shows that mode choice is 3, which is "satisfied." The total SIM LG leader respondents of 85.5% chose "satisfied" or "extremely satisfied."

Table 304: Result of the SIM LG leaders' satisfaction level rating median calculation

```
. tabstat q9, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q9 |        76         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is “satisfied.” This means at least 50% of the SIM LG leader respondents are in the “satisfied” or “extremely satisfied” category looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 305: Result of the SIM LG leaders’ measure of consensus on satisfaction level

```
. cns q9 , min(1) max(4)

Consensus Measure for q9
Cns(X) = .73416144
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the satisfaction level of SIM LG leaders, it is 0.7342.

Inferential Analysis – Statistical Significance Testing

Table 306: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q9 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       65      2574.5        1463
    Negative |       11       351.5        1463
        Zero |        0           0           0
-------------+---------------------------------
         All |       76        2926        2926

Unadjusted variance    37306.50
Adjustment for ties    -3823.75
Adjustment for zeros       0.00
                     ----------
Adjusted variance      33482.75

H0: q9 = 2.5
         z =  6.074
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 85.5% of SIM LG leaders surveyed think that SIM programme was satisfactory. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM LG LEADER population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "dissatisfied" and 3 = "satisfied."

Ho: Our null hypothesis is that the SIM LG leader population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM LG leader population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.074, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 6.074 and our sample size for SIM LG leaders is 76. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 6.074/SQRT(76) = 0.70. This, according to Bartz (1999), is strong effect size.

Evidence on SIM LG leaders' Satisfaction Level

There is statistically significant evidence (p = 0.0000) that 85.5% of SIM LG leaders are satisfied with the MOE's SIM programme during COVID-19 pandemic as an Education in Emergency intervention. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.074, p = 0.0000, with a strong effect size (*r* = 0.70).

Analyzing LG leaders' Perception on Implementation of SIM

The Ministry of Education was interested to know how effectively implementation of SIM programme was carried out in the perception of LG leaders. To investigate this, ***Figure 77***, which is visualization of survey data, shows the results of implementation effectiveness perception of LG leaders from the SIM survey.

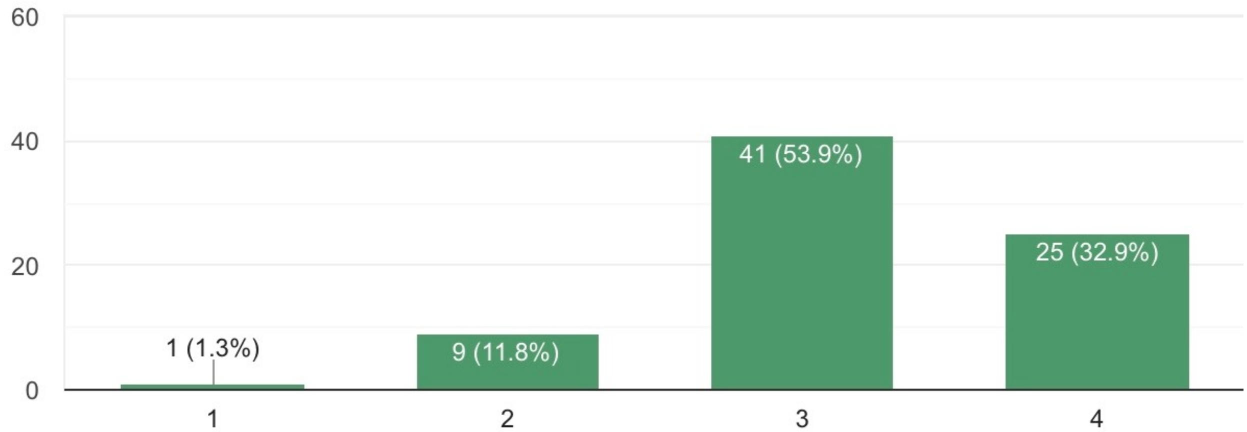


Figure 77*: Results of "Rate how effectively has the SIM been implemented" where 1 = Not effective, 2 = Slightly effective, 3 = Effective, and 4 = Very effective*

As can be seen in ***Figure 77*** the 86.8% of the SIM LG leader respondents rated that the SIM programme implementation was "effective" or "very effective."

Descriptive Analysis – Measure of Central Tendency

Table 307: Results of the SIM LG leaders' SIM implementation rating frequency distribution

```
. tabulate q8

         q8 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          1        1.32        1.32
          2 |          9       11.84       13.16
          3 |         41       53.95       67.11
          4 |         25       32.89      100.00
------------+-----------------------------------
      Total |         76      100.00
```

From the frequency ***Table 307*** above, it shows that mode choice is 3, which is "effective." The total SIM LG leader respondents of 86.8% chose "effective" or "very effective" in their perception on implementation effectiveness of the SIM.

Table 308: Result of the SIM LG leaders' SIM implementation rating median calculation

```
. tabstat q8, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q8 |        76         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM LG leader respondents believe that SIM implementation was "effective" or "very effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 309: Result of the SIM LG leaders' measure of consensus on SIM implementation

```
. cns q8 , min(1) max(4)

Consensus Measure for q8
Cns(X) = .68957117
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the implementation effectiveness opinion of SIM LG leaders, it is 0.6896.

Inferential Analysis – Statistical Significance Testing

Table 310: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       66        2633        1463
    Negative |       10         293        1463
        Zero |        0           0           0
-------------+---------------------------------
         All |       76        2926        2926

Unadjusted variance    37306.50
Adjustment for ties    -2968.75
Adjustment for zeros       0.00
                     ----------
Adjusted variance      34337.75

H0: q8 = 2.5
         z =   6.314
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 86.8% of SIM LG leaders surveyed think that SIM programme was effectively implemented. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM LG leader population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly effective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM LG leader population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM LG leader population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.314, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 6.314 and our sample size for SIM LG leaders is 76. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 6.314/SQRT(76) = 0.72. This, according to Bartz (1999), is strong effect size.

Evidence on SIM LG leaders' Perception on Implementation Effectiveness of SIM

There is statistically significant evidence (p = 0.0000) that 86.8% of SIM LG leaders believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 6.314, p = 0.0000, with a strong effect size (r = 0.72).

Analyzing LG Leaders' Perception on Usefulness of SIM

The Ministry of Education was interested to know how useful was SIM programme in the perception of LG leaders. To investigate this, ***Figure 78***, which is visualization of survey data, shows the results of SIM usefulness perception of LG leaders from the SIM survey.

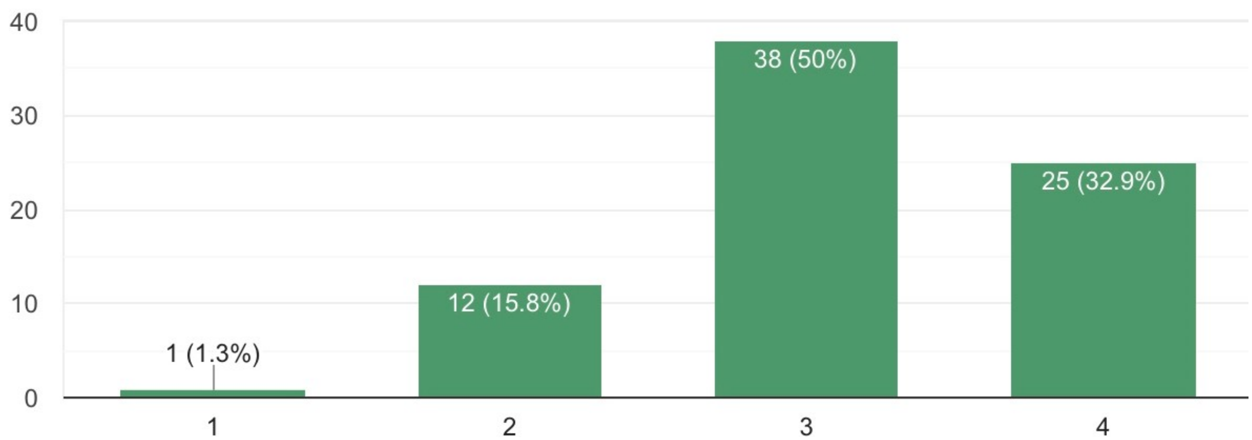


***Figure 78**: Results of "Rate how useful was SIM" where 1 = Not useful, 2 = Slightly useful, 3 = Useful, and 4 = Very useful*

As can be seen in ***Figure 78*** the 82.9% of the SIM LG respondents rated that the SIM programme "useful" or "very useful."

Descriptive Analysis – Measure of Central Tendency

Table 311: Results of the SIM LG leaders' SIM usefulness rating frequency distribution

```
. tabulate q7

         q7 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          1        1.32        1.32
          2 |         12       15.79       17.11
          3 |         38       50.00       67.11
          4 |         25       32.89      100.00
------------+-----------------------------------
      Total |         76      100.00
```

From the frequency ***Table 311*** above, it shows that mode choice is 3, which is "useful." The total SIM LG respondents of 82.9% chose "useful" or "very useful" in their perception on usefulness of SIM.

Table 312: Result of the SIM LG leaders' SIM usefulness rating median calculation

```
. tabstat q7, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q7 |        76         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "useful." This means at least 50% of the SIM LG respondents believe that SIM was "useful" or "very useful" looking at the median score of 3.

Descriptive Analysis - Measure of Dispersion

Table 313: Result of the SIM LG leaders' measure of consensus on SIM usefulness

```
. cns q7 , min(1) max(4)

Consensus Measure for q7
Cns(X) = .67176968
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the SIM usefulness opinion of SIM LG leaders, it is 0.6718.

Inferential Analysis – Statistical Significance Testing

Table 314: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q7 = 2.5

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |       63      2556.5        1463
    Negative |       13       369.5        1463
        Zero |        0           0           0
-------------+---------------------------------
         All |       76        2926        2926

Unadjusted variance    37306.50
Adjustment for ties    -2968.75
Adjustment for zeros       0.00
                     ----------
Adjusted variance      34337.75

H0: q7 = 2.5
         z =  5.901
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 82.9% of SIM LG leaders surveyed think that SIM programme was useful. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM LG population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly useful" and 3 = "useful."

Ho: Our null hypothesis is that the SIM LG population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM LG population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 5.901, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 5.901 and our sample size for SIM LG leaders is 76. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 5.901/SQRT(76) = 0.68. This, according to Bartz (1999), is strong effect size.

Evidence on SIM LG leaders' Perception on Usefulness of SIM

There is statistically significant evidence ($p = 0.0000$) that 82.9% of SIM LG leaders believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, $Z = 5.901$, $p = 0.0000$, with a strong effect size ($r = 0.68$).

Analyzing LG Leaders' Perception on Overall Presentation of SIM

The Ministry of Education was interested to know how attractive was overall presentation of SIM booklets in the perception of LG leaders. To investigate this, ***Figure 79***, which is visualization of survey data, shows the results of overall presentation of SIM booklets in the perception of LG leaders.

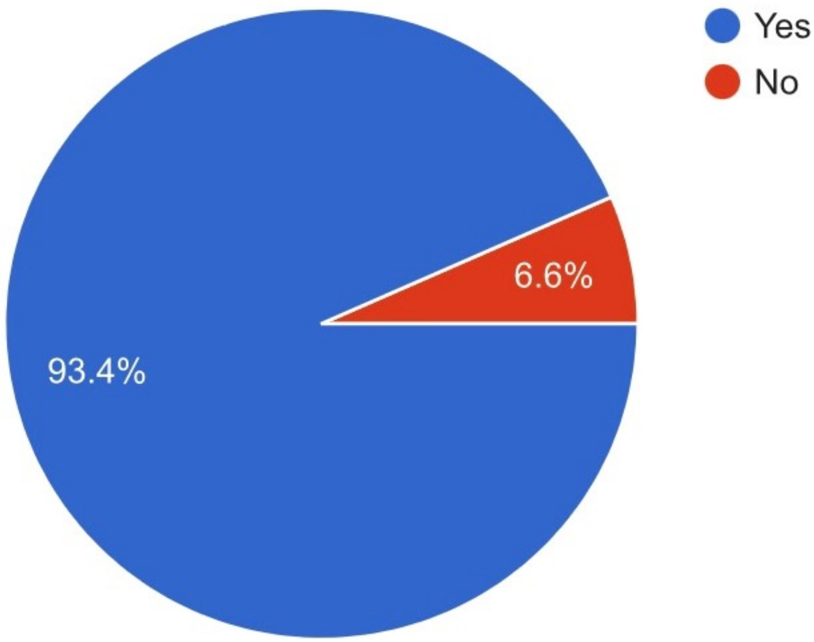


Figure 79*: Results of "Is overall presentation of SIM attractive?"*

As can be seen in ***Figure 79*** the 93.4% of the SIM LG respondents rated that overall presentation of SIM booklets is attractive.

Inferential Analysis – Statistical Significance Testing

Table 315: Results of Binomial Test on LG leaders' Perception on Overall Presentation of SIM

```
. bitest q6a = 0.86

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q6a |         76          71        65.36     0.86000      0.93421

  Pr(k >= 71)            = 0.035814  (one-sided test)
  Pr(k <= 71)            = 0.986400  (one-sided test)
  Pr(k <= 59 or k >= 71) = 0.067806  (two-sided test)
```

A binomial test indicated that the percentage of SIM LG leaders who believe overall presentation of SIM is attractive (N_{Yes} = 71, 93.4%), was statistically significantly greater than the population hypothesized value of 86%, p = 0.035814.

Evidence on SIM LG leaders' Perception on Overall Presentation of SIM

There is statistically significant evidence (p = 0.035814) that at least 86% of SIM LG leaders believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM LG leaders who believe overall presentation of SIM is attractive (N_{Yes} = 71, 93.4%), was statistically significantly greater than the population hypothesized value of 86%, p = 0.035814.

Analyzing LG leaders' Perception on Whether They Delivered SIMs

The Ministry of Education was interested to know whether the gewog office provided support in delivering the SIMs. To investigate this, ***Figure 80***, which is visualization of survey data, shows the results of whether gewog offices provided support in delivering SIMs or not, in the perception of LG leaders.

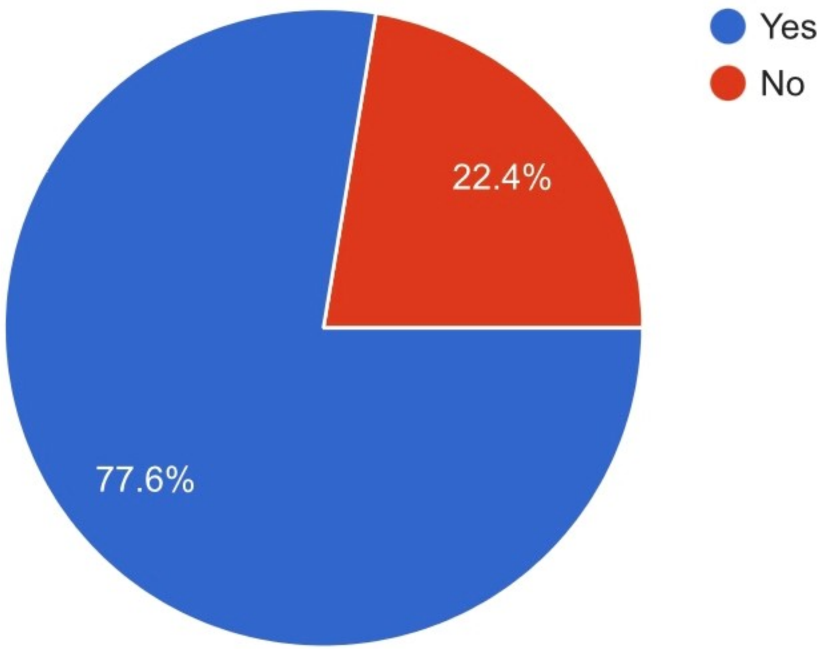


Figure 80: *Results of "Did your gewog office provide support in delivering the SIMs?"*

As can be seen in ***Figure 80*** the 77.6% of the SIM LG respondents said that their gewog office provided support in delivering the SIMs.

Inferential Analysis – Statistical Significance Testing

Table 316: Results of Binomial Test on Whether Gewog Offices Provided Support for the SIMs

```
. bitest q5a = 0.67

Binomial probability test

Variable |         N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q5a |        76          59        50.92     0.67000      0.77632

  Pr(k >= 59)            = 0.029282  (one-sided test)
  Pr(k <= 59)            = 0.984379  (one-sided test)
  Pr(k <= 42 or k >= 59) = 0.051019  (two-sided test)
```

A binomial test indicated that the percentage of SIM LG leaders who believe their offices supported in delivering the SIMs (N_{Yes} = 59, 77.6%), was statistically significantly greater than the population hypothesized value of 67%, $p = 0.029282$.

Evidence on SIM LG leaders' Perception on Whether Gewog Offices Delivered the SIMs

There is statistically significant evidence (p = 0.029282) that at least 67% of SIM LG leaders believe their offices delivered the SIMs. A binomial test indicated that the percentage of SIM LG leaders who believe their offices supported in delivering the SIMs (N_{Yes} = 59, 77.6%), was statistically significantly greater than the population hypothesized value of 67%, $p = 0.029282$.

Analyzing LG Leaders' Perception on Whether SIM Reached the Identified Students

The Ministry of Education was interested to know whether the SIM reached the identified students. To investigate this, ***Figure 81***, which is visualization of survey data, shows the results of whether SIM reached the identified students or not, in the perception of LG leaders.

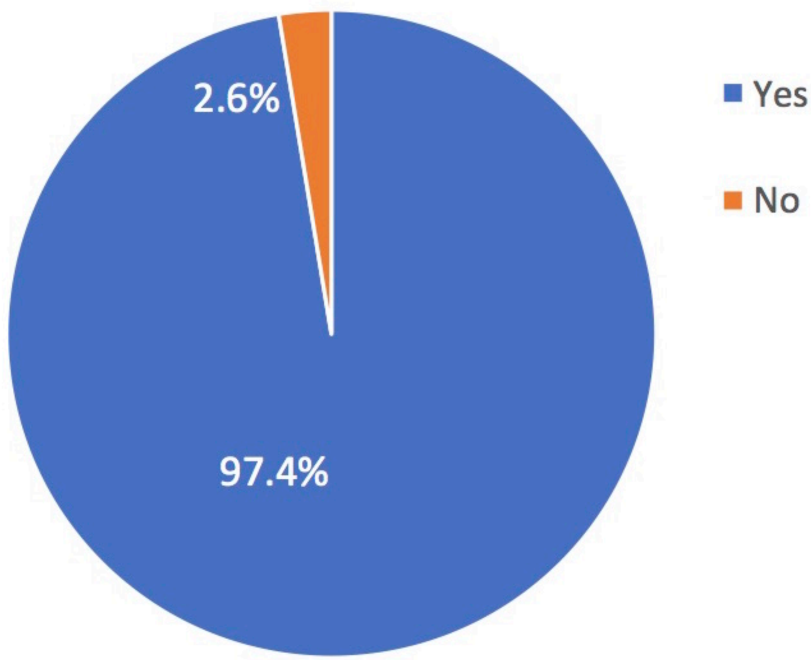


Figure 81: *Results of "Has the SIM reached the identified students?"*

As can be seen in ***Figure 81*** the 97.4% of the SIM LG respondents said that the SIM has reached the identified students.

Inferential Analysis – Statistical Significance Testing

Table 317: Results of Binomial Test on Whether SIM Reached the Identified Students

```
. bitest q3a = 0.91

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q3a |       76          74        69.16     0.91000      0.97368

  Pr(k >= 74)            = 0.028065  (one-sided test)
  Pr(k <= 74)            = 0.993432  (one-sided test)
  Pr(k <= 64 or k >= 74) = 0.066672  (two-sided test)
```

A binomial test indicated that the percentage of SIM LG leaders who believe the SIM has reached the identified students (N_{Yes} = 74, 97.4%), was statistically significantly greater than the population hypothesized value of 91%, p = 0.028065.

Evidence on SIM LG leaders' Perception on Whether SIM Reached the Identified Students

There is statistically significant evidence (p = 0.028065) that at least 91% of SIM LG leaders believe SIM has reached the identified students. A binomial test indicated that the percentage of SIM LG leaders who believe the SIM has reached the identified students (N_{Yes} = 74, 97.4%), was statistically significantly greater than the population hypothesized value of 91%, p = 0.028065.

Analyzing LG leaders' Perception on Whether SIM Reached Other Needy Students

The Ministry of Education was interested to know whether the SIM reached other needy students beyond the identified students. To investigate this, ***Figure 82***, which is visualization of survey data, shows the results of whether SIM reached other needy students beyond the identified students or not, in the perception of LG leaders.

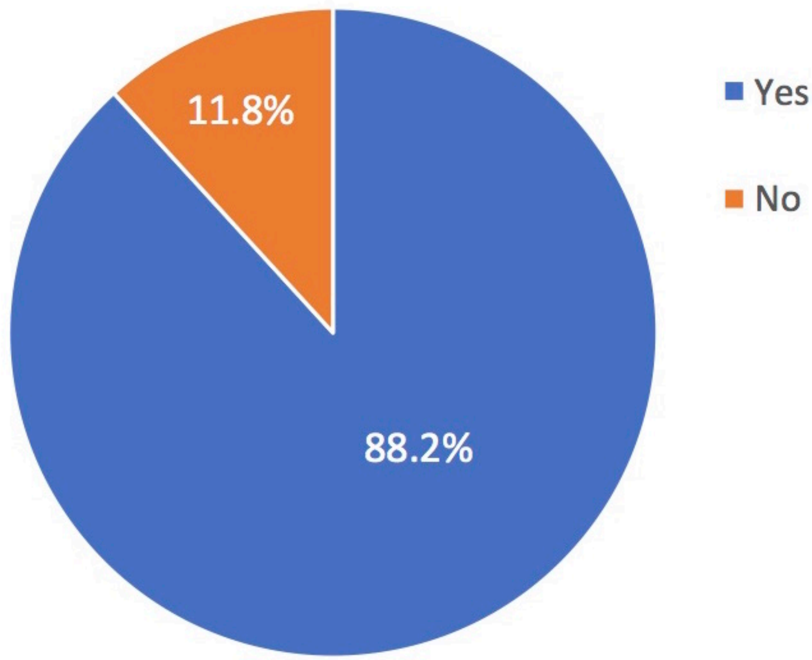


***Figure 82**: Results of "Has the SIM reached other needy students beyond the identified students?"*

As can be seen in ***Figure 82*** the 88.2% of the SIM LG respondents said that the SIM has reached other needy students.

Inferential Analysis – Statistical Significance Testing

Table 318: Results of Binomial Test on Whether SIM Reached Other Needy Students

```
. bitest q4a = 0.79

Binomial probability test

Variable |          N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q4a |         76           67        60.04     0.79000      0.88158

  Pr(k >= 67)            = 0.028670  (one-sided test)
  Pr(k <= 67)            = 0.986987  (one-sided test)
  Pr(k <= 52 or k >= 67) = 0.048980  (two-sided test)
```

A binomial test indicated that the percentage of SIM LG leaders who believe the SIM has reached other needy students (N_{Yes} = 67, 88.2%), was statistically significantly greater than the population hypothesized value of 79%, p = 0.028670.

Evidence on SIM LG leaders' Perception on Whether SIM Reached Other Needy Students

There is statistically significant evidence (p = 0.028670) that at least 79% of SIM LG leaders believe SIM has reached other needy students. A binomial test indicated that the percentage of SIM LG leaders who believe the SIM has reached other needy students (N_{Yes} = 67, 88.2%), was statistically significantly greater than the population hypothesized value of 79%, p = 0.028670.

PART VI: SIM Parents

Demographic Characteristics of SIM Parent Respondents

The age characteristics of the SIM parent respondents are summarized in ***Table 319***. The age of the SIM parent respondents ranged from 19 to 72 years (M = 37.93, SD = 8.45).

Table 319: Results of age characteristics of SIM parent respondents

Variable	Obs	Mean	Std. dev.	Min	Max
age	374	37.92513	8.44917	19	72

Similarly, among the 374 SIM principal respondents, 166 (44.4%) were males and 208 (55.6%) were females as shown in ***Figure 83***.

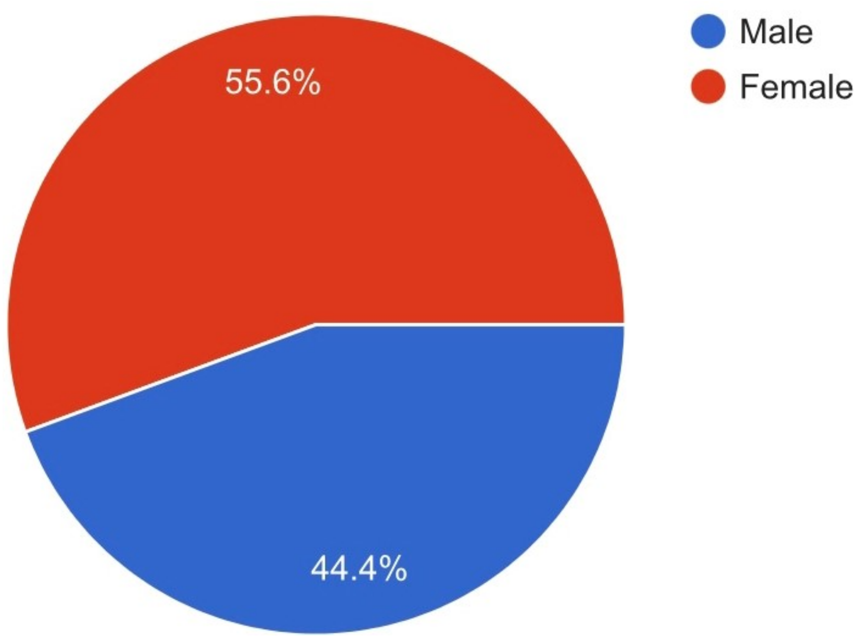


Figure 83: *Gender of SIM parent respondents*

Among the 374 SIM parent respondents, we got data representation from all types of schools such as HSS (15.2%), MSS (20.6%), LSS (11.5%), PS (40.4%), and ECR (12.3%) as shown in ***Figure 84***.

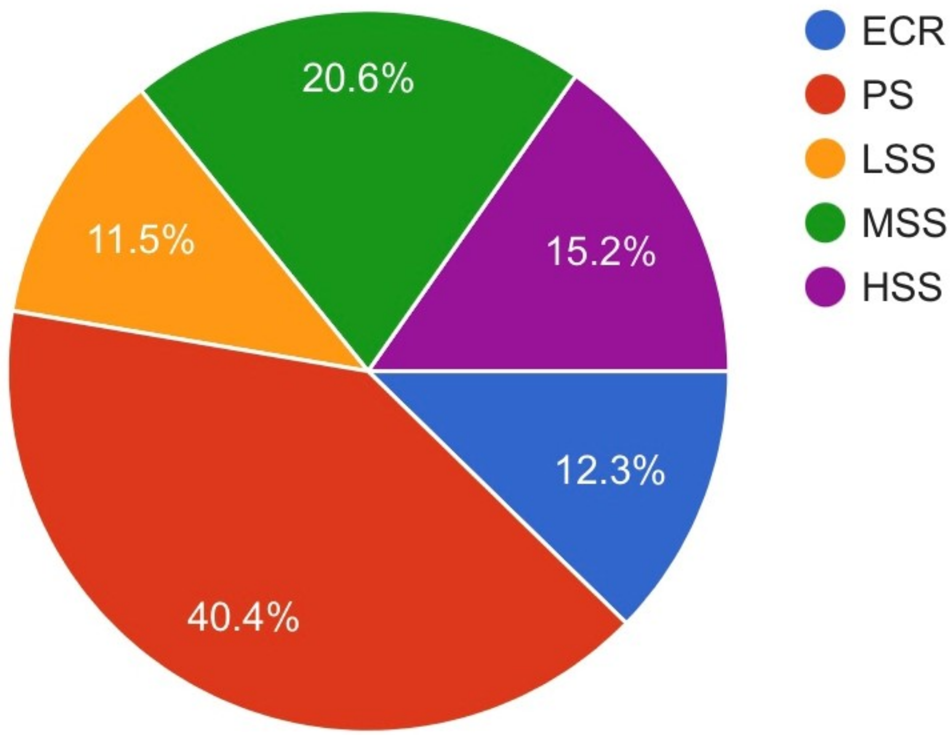


Figure 84*: School types of SIM parent respondents*

We also included question on special education needs (SEN) students. Among the 374 SIM parent respondents, 34 (9.1%) said their children are SEN students and 340 (90.9%) said their children are not SEN students as shown in ***Figure 85***.

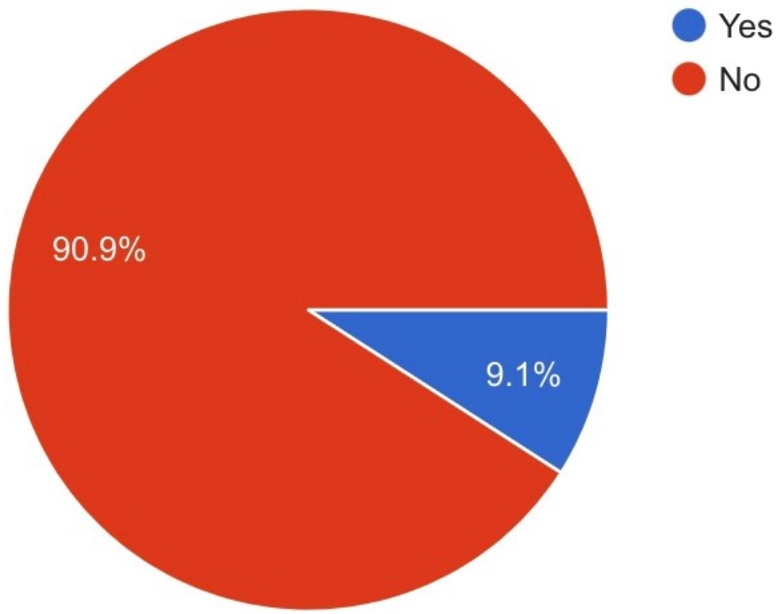


Figure 85*: Results of "Is your child a special education needs (SEN) student?"*

Effectiveness of SIM Programme

Analyzing Parents’ Perception on Implementation of SIM

The Ministry of Education was interested to know how effectively implementation of SIM programme was carried out in the perception of parents. To investigate this, ***Figure 86***, which is visualization of survey data, shows the results of implementation effectiveness perception of parents from the SIM survey.

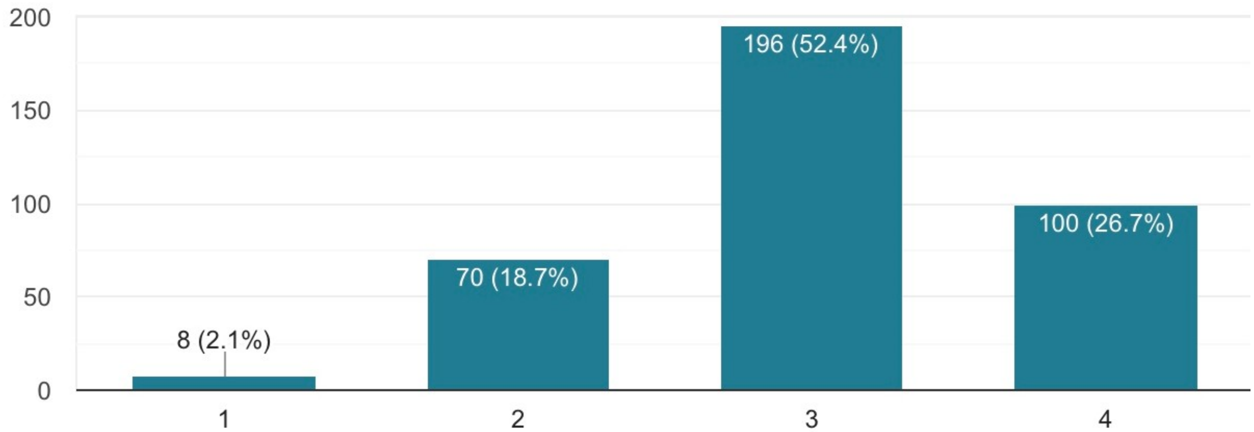


Figure 86*: Results of “Rate how effectively has the SIM been implemented” where 1 = Not effective, 2 = Slightly effective, 3 = Effective, and 4 = Very effective*

As can be seen in ***Figure 86*** the 79.1% of the SIM parent respondents rated that the SIM programme implementation was “effective” or “very effective.”

Descriptive Analysis – Measure of Central Tendency

Table 320: Results of the SIM Parents’ SIM implementation rating frequency distribution

```
. tabulate q12

        q12 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          8        2.14        2.14
          2 |         70       18.72       20.86
          3 |        196       52.41       73.26
          4 |        100       26.74      100.00
------------+-----------------------------------
      Total |        374      100.00
```

From the frequency ***Table 320*** above, it shows that mode choice is 3, which is “effective.” The total SIM parent respondents of 79.1% chose “effective” or “very effective” in their perception on implementation effectiveness of the SIM.

Table 321: Result of the SIM Parents' SIM implementation rating median calculation

```
. tabstat q12, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
         q12 |       374         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "effective." This means at least 50% of the SIM parent respondents believe that SIM implementation was "effective" or "very effective" looking at the median score rating of 3.

Descriptive Analysis - Measure of Dispersion

Table 322: Result of the SIM Parents' measure of consensus on SIM implementation

```
. cns q12 , min(1) max(4)

Consensus Measure for q12
Cns(X) = .69158569
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the implementation effectiveness opinion of SIM parents, it is 0.6916.

Inferential Analysis – Statistical Significance Testing

Table 323: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q12 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      296       58216     35062.5
    Negative |       78       11909     35062.5
        Zero |        0           0           0
-------------+---------------------------------
         All |      374       70125       70125

Unadjusted variance  4376968.75
Adjustment for ties  -418342.38
Adjustment for zeros       0.00
                     ----------
Adjusted variance    3958626.38

H0: q12 = 2.5
         z = 11.637
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 79.1% of SIM parents surveyed think that SIM programme was effectively implemented. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM parent population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly effective" and 3 = "effective."

Ho: Our null hypothesis is that the SIM parent population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM parent population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 11.637, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 11.637 and our sample size for SIM Parents is 374. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 11.637/SQRT(374) = 0.60. This, according to Bartz (1999), is moderate effect size.

Evidence on SIM Parents' Perception on Implementation Effectiveness of SIM

There is statistically significant evidence (p = 0.0000) that 79.1% of SIM parents believe the SIM programme implementation was effective. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 11.637, p = 0.0000, with a moderate effect size (*r* = 0.60).

Analyzing Parents' Perception on Usefulness of SIM

The Ministry of Education was interested to know how useful was SIM programme in the perception of parents. To investigate this, ***Figure 87***, which is visualization of survey data, shows the results of SIM usefulness perception of parents from the SIM survey.

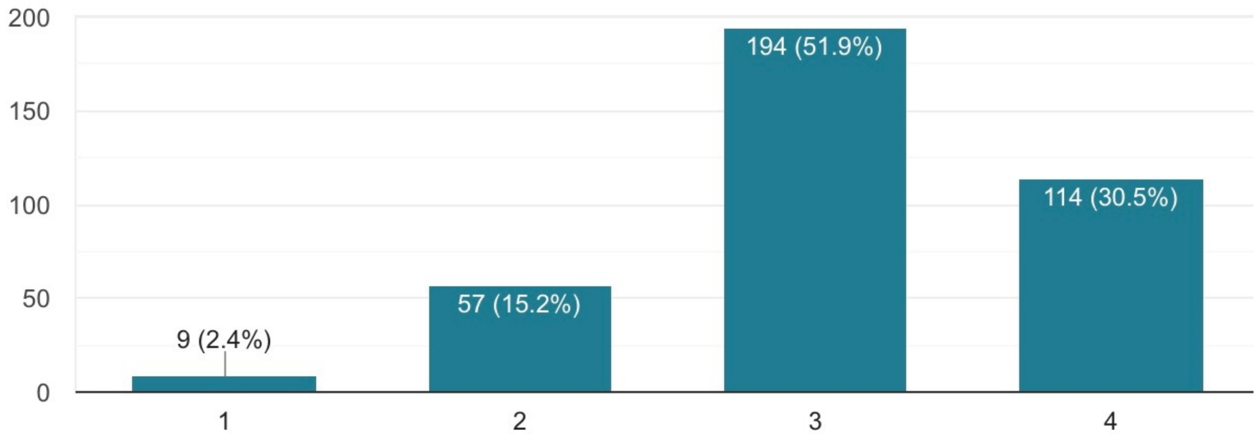


***Figure 87**: Results of "Rate how useful was SIM" where 1 = Not useful, 2 = Slightly useful, 3 = Useful, and 4 = Very useful*

As can be seen in ***Figure 87*** the 82.4% of the SIM parent respondents rated that the SIM programme "useful" or "very useful."

Descriptive Analysis – Measure of Central Tendency

Table 324: Results of the SIM Parents' SIM usefulness rating frequency distribution

```
. tabulate q8

         q8 |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |          9        2.41        2.41
          2 |         57       15.24       17.65
          3 |        194       51.87       69.52
          4 |        114       30.48      100.00
------------+-----------------------------------
      Total |        374      100.00
```

From the frequency ***Table 324*** above, it shows that mode choice is 3, which is "useful." The total SIM parent respondents of 82.4% chose "useful" or "very useful" in their perception on usefulness of SIM.

Table 325: Result of the SIM parents' SIM usefulness rating median calculation

```
. tabstat q8, stat(count p50 min max)

    Variable |         N       p50       Min       Max
-------------+----------------------------------------
          q8 |       374         3         1         4
------------------------------------------------------
```

The calculated sample median = 3, which is "useful." This means at least 50% of the SIM parent respondents believe that SIM was "useful" or "very useful" looking at the median score of 3.

Descriptive Analysis - Measure of Dispersion

Table 326: Result of the SIM Parents' measure of consensus on SIM usefulness

```
. cns q8 , min(1) max(4)

Consensus Measure for q8
Cns(X) = .67468154
```

The measure of consensus (Tastle & Wierman, 2007), which ranges from 0 to 1 where 0 is complete lack of consensus and 1 is complete consensus of opinion, shows that for the SIM usefulness opinion of SIM parents, it is 0.6747.

Inferential Analysis – Statistical Significance Testing

Table 327: Results of One-Sample Wilcoxon Signed Rank Test

```
. signrank q8 = 2.5, exact

Wilcoxon signed-rank test

        Sign |      Obs   Sum ranks    Expected
-------------+---------------------------------
    Positive |      308       60126     35062.5
    Negative |       66        9999     35062.5
        Zero |        0           0           0
-------------+---------------------------------
         All |      374       70125       70125

Unadjusted variance  4376968.75
Adjustment for ties  -368203.00
Adjustment for zeros      0.00
                     ----------
Adjusted variance    4008765.75

H0: q8 = 2.5
         z = 12.518
Prob > |z| = 0.0000
Exact prob = 0.0000
```

We have seen that the 82.4% of SIM parents surveyed think that SIM programme was useful. However, that was just based on our sample from the SIM survey. We need to test whether this would be the case in the SIM parent population too. In other words, we have to test whether the true median score in the population would be statistically significantly different from 2.5 since 2 = "slightly useful" and 3 = "useful."

Ho: Our null hypothesis is that the SIM parent population's true median is 2.5.

Ha: Our alternative hypothesis is that the SIM parent population's true median is significantly different from 2.5.

One-sample Wilcoxon signed rank test would show us how likely to have result as in our survey sample or even more extreme if the true median in the population is 2.5. Since our p-value is 0.0000, which is very low or significantly less than alpha = 0.05, we can conclude that the true population median is statistically significantly different from 2.5. In short, our one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 12.518, p = 0.0000. The positive z-score shows that the population median is above the hypothesized median of 2.5.

Effect Size

The test statistic is Z = 12.518 and our sample size for SIM Parents is 374. Therefore, the effect size (Rosenthal, 1991), r = test statistic/square root of sample size, which is 12.518/SQRT(374) = 0.65. This, according to Bartz (1999), is strong effect size.

Evidence on SIM Parents' Perception on Usefulness of SIM

There is statistically significant evidence (p = 0.0000) that 82.4% of SIM parents believe the SIM programme was useful. In particular, one-sample Wilcoxon signed rank test indicated that the population median was significantly different from 2.5, Z = 12.518, p = 0.0000, with a strong effect size (r = 0.65).

Analyzing Parents' Perception on Overall Presentation of SIM

The Ministry of Education was interested to know how attractive was overall presentation of SIM booklets in the perception of parents. To investigate this, ***Figure 88***, which is visualization of survey data, shows the results of overall presentation of SIM booklets in the perception of parents.

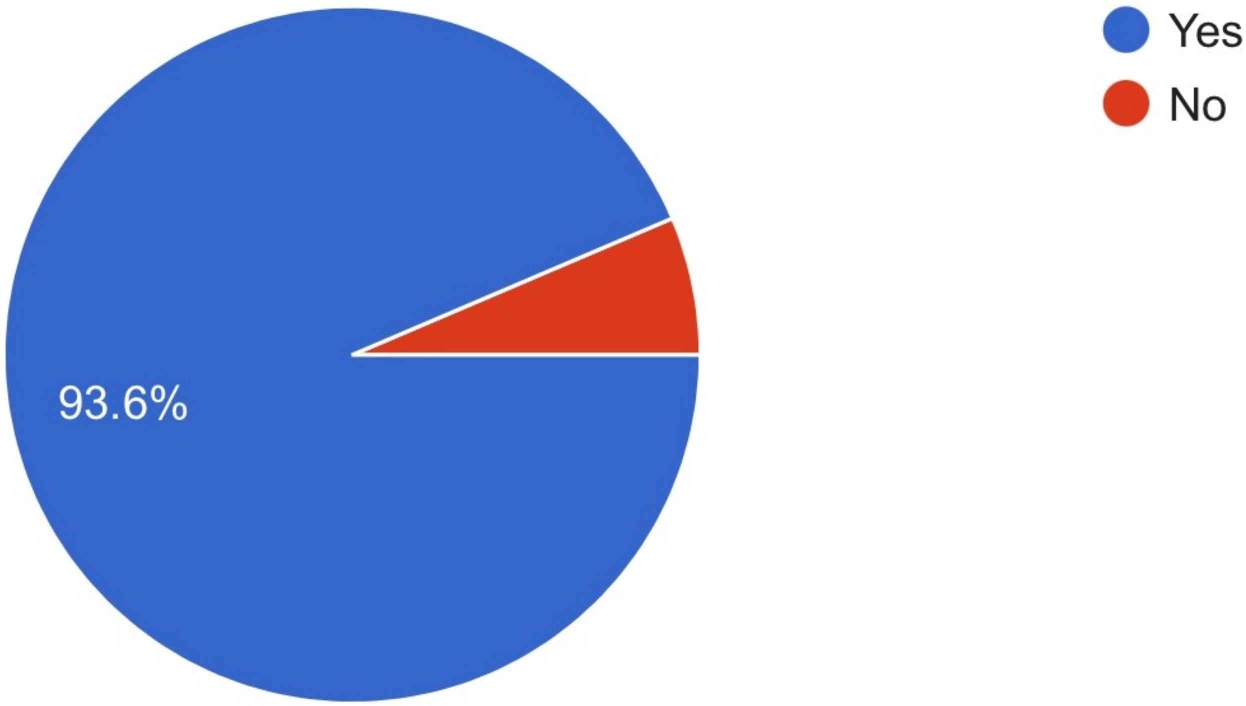


Figure 88: *Results of "Is overall presentation of SIM attractive?"*

As can be seen in ***Figure 88*** the 93.6% of the SIM parent respondents rated that overall presentation of SIM booklets is attractive.

Inferential Analysis – Statistical Significance Testing

Table 328: Results of Binomial Test on Parents' Perception on Overall Presentation of SIM

```
. bitest q11a = 0.90

Binomial probability test

Variable |         N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
    q11a |       374          350        336.6     0.90000      0.93583

  Pr(k >= 350)            = 0.009820  (one-sided test)
  Pr(k <= 350)            = 0.994288  (one-sided test)
  Pr(k <= 322 or k >= 350) = 0.019607  (two-sided test)
```

A binomial test indicated that the percentage of SIM parents who believe overall presentation of SIM is attractive (N_{Yes} = 350, 93.6%), was statistically significantly greater than the population hypothesized value of 90%, p = 0.009820.

Evidence on SIM Parents' Perception on Overall Presentation of SIM

There is statistically significant evidence (p = 0.009820) that at least 90% of SIM parents believe overall presentation of SIM is attractive. A binomial test indicated that the percentage of SIM parents who believe overall presentation of SIM is attractive (N_{Yes} = 350, 93.6%), was statistically significantly greater than the population hypothesized value of 90%, p = 0.009820.

Analyzing Parents' Perception on Schools' Support Extended to SIM Students

The Ministry of Education was interested to know about help extended to SIM students in the perception of parents. To investigate this, ***Figure 89***, which is visualization of survey data, shows the results of support extended to SIM students in the perception of parents.

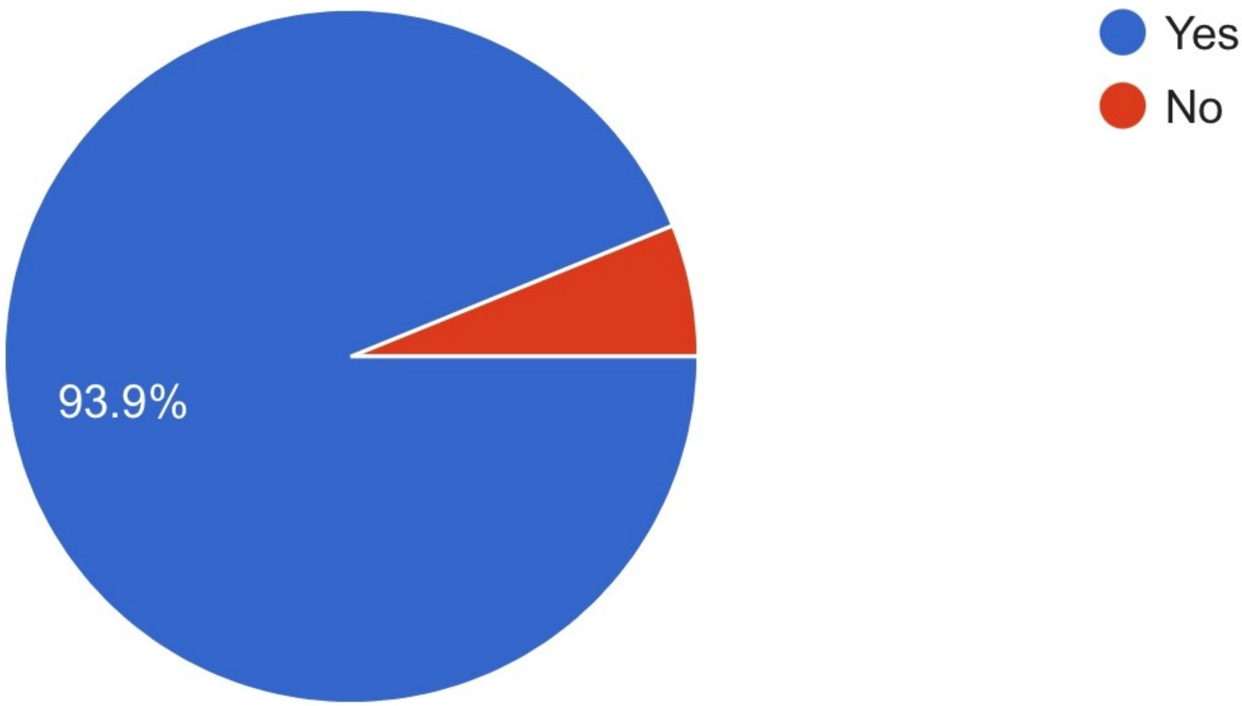


Figure 89*: Results of "Did the school offer any help to your child?"*

As can be seen in ***Figure 89*** the 93.9% of the SIM parent respondents said that their schools offered help to their children.

Inferential Analysis – Statistical Significance Testing

Table 329: Results of Binomial Test on Help Offered to SIM Children

```
. bitest q10a = 0.91

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
    q10a |      374          351       340.34     0.91000      0.93850

  Pr(k >= 351)            = 0.028362  (one-sided test)
  Pr(k <= 351)            = 0.982472  (one-sided test)
  Pr(k <= 329 or k >= 351) = 0.057184  (two-sided test)
```

A binomial test indicated that the percentage of SIM parents who believe their schools offered help to their SIM children (N_{Yes} = 351, 93.9%), was statistically significantly greater than the population hypothesized value of 91%, $p = 0.028362$.

Evidence on SIM Parents' Perception on Help Offered to SIM Children

There is statistically significant evidence (p = 0.028362) that at least 91% of SIM parents believe the schools offered help to their SIM children. A binomial test indicated that the percentage of SIM parents who believe their schools offered help to their SIM children (N_{Yes} = 351, 93.9%), was statistically significantly greater than the population hypothesized value of 91%, $p = 0.028362$.

Analyzing Parents' Perception on Help Sought by SIM Children

The Ministry of Education was interested to know about help sought by SIM students in the perception of parents. To investigate this, ***Figure 90***, which is visualization of survey data, shows the results of help sought by SIM children in the perception of parents.

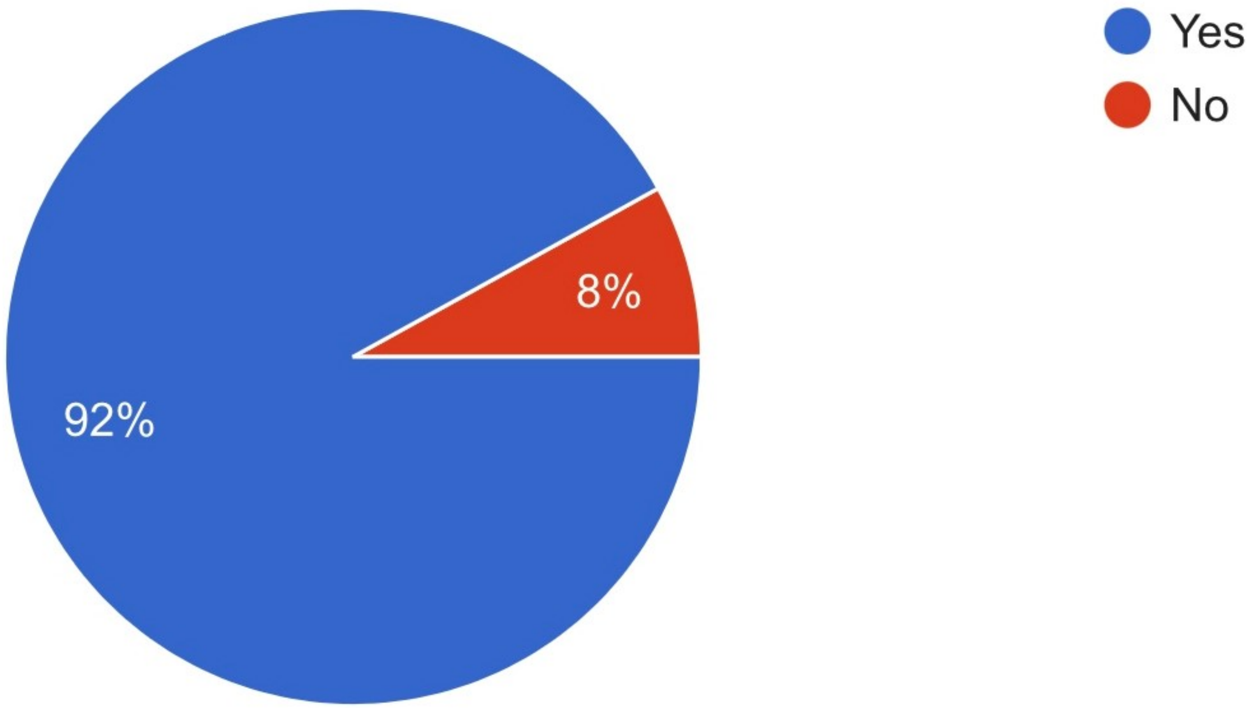


Figure 90: *Results of "Did your child seek help from anyone to understand the lessons?"*

As can be seen in ***Figure 90*** the 92.0% of the SIM parent respondents said that their children sought help to understand SIM lessons.

Inferential Analysis – Statistical Significance Testing

Table 330: Results of Binomial Test on Help Sought by SIM Children

```
. bitest q9a = 0.89

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q9a |      374          344       332.86     0.89000      0.91979

  Pr(k >= 344)            = 0.035098  (one-sided test)
  Pr(k <= 344)            = 0.976800  (one-sided test)
  Pr(k <= 321 or k >= 344) = 0.068726  (two-sided test)
```

A binomial test indicated that the percentage of SIM parents who believe their children sought help to understand SIM lessons (N_{Yes} = 344, 92.0%), was statistically significantly greater than the population hypothesized value of 89%, p = 0.035098.

Evidence on SIM Parents' Perception on Help Sought by SIM Children

There is statistically significant evidence (p = 0.035098) that at least 89% of SIM parents believe their children sought help regarding SIM lessons. A binomial test indicated that the percentage of SIM parents who believe their children sought help to understand SIM lessons (N_{Yes} = 344, 92.0%), was statistically significantly greater than the population hypothesized value of 89%, p = 0.035098.

Analyzing Parents' Perception on Whether Their Children Use SIM for Self-Learning

The Ministry of Education was interested to know whether SIM children use SIM for self-learning. To investigate this, ***Figure 91***, which is visualization of survey data, shows the results of whether SIM children use SIM for self-learning, in the perception of parents.

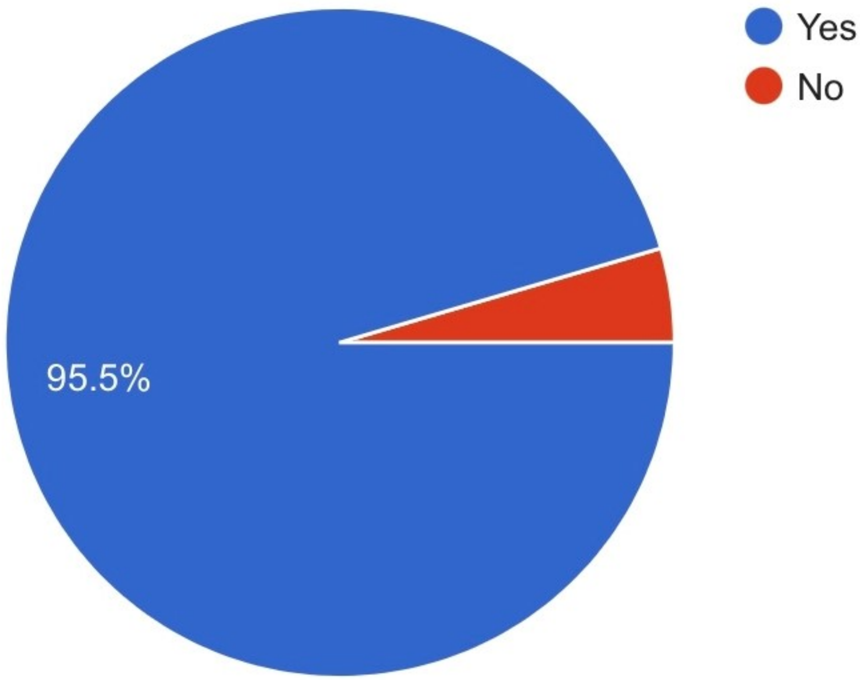


Figure 91*: Results of "Does your child use SIM for self-learning?"*

As can be seen in ***Figure 91*** the 95.5% of the SIM parent respondents said that their children use SIM for self-learning.

Inferential Analysis – Statistical Significance Testing

Table 331: Results of Binomial Test on Whether SIM Children Use SIM for Self-Learning

```
. bitest q7a = 0.93

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q7a |      374          357       347.82     0.93000      0.95455

  Pr(k >= 357)           = 0.033387  (one-sided test)
  Pr(k <= 357)           = 0.980497  (one-sided test)
  Pr(k <= 338 or k >= 357) = 0.067331  (two-sided test)
```

A binomial test indicated that the percentage of SIM parents who believe their children used SIM for self-learning (N_{Yes} = 357, 95.5%), was statistically significantly greater than the population hypothesized value of 93%, p = 0.033387.

Evidence on SIM Parents' Perception on Their Children's Use SIM for Self-Learning

There is statistically significant evidence (p = 0.033387) that at least 93% of SIM parents believe their children used SIM for self-learning. A binomial test indicated that the percentage of SIM parents who believe their children used SIM for self-learning (N_{Yes} = 357, 95.5%), was statistically significantly greater than the population hypothesized value of 93%, p = 0.033387.

Analyzing Parents' Perception on Whether Their Children Received SIM

The Ministry of Education was interested to know whether SIM children received SIM. To investigate this, ***Figure 92***, which is visualization of survey data, shows the results of whether SIM children received SIM, in the perception of parents.

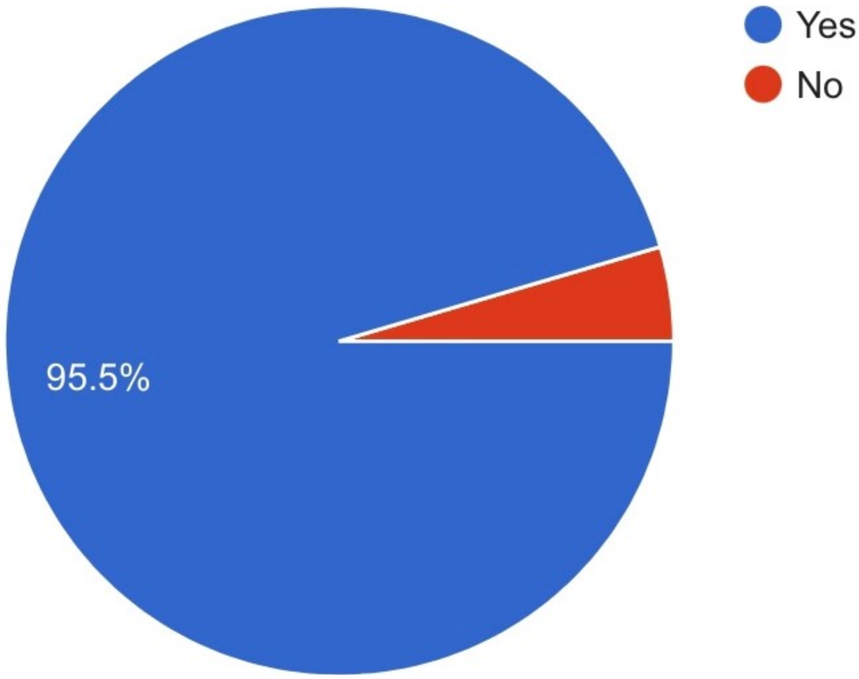


***Figure 92**: Results of "Did your child receive SIM?"*

As can be seen in ***Figure 92*** the 95.5% of the SIM parent respondents said that their children received SIM.

Inferential Analysis – Statistical Significance Testing

Table 332: Results of Binomial Test on Whether SIM Children Received SIM

```
. bitest q6a = 0.93

Binomial probability test

Variable |        N   Observed k   Expected k   Assumed p   Observed p
-------------+------------------------------------------------------------
     q6a |      374          357       347.82     0.93000      0.95455

  Pr(k >= 357)            = 0.033387  (one-sided test)
  Pr(k <= 357)            = 0.980497  (one-sided test)
  Pr(k <= 338 or k >= 357) = 0.067331  (two-sided test)
```

A binomial test indicated that the percentage of SIM parents who believe their children received SIM (N_{Yes} = 357, 95.5%), was statistically significantly greater than the population hypothesized value of 93%, $p = 0.033387$.

Evidence on SIM Parents' Perception on Whether Their Children Received SIM

There is statistically significant evidence (p = 0.033387) that at least 93% of SIM parents believe their children received SIM. A binomial test indicated that the percentage of SIM parents who believe their children received SIM (N_{Yes} = 357, 95.5%), was statistically significantly greater than the population hypothesized value of 93%, $p = 0.033387$.

Conclusion

This nationwide SIM assessment study showed that SIM learning is a valuable method of teaching students as an Education in Emergency intervention. In the opinion of the respondents in our survey, there is enough evidence that SIM programme is satisfactory and is accepted. The survey respondents also rated the overall presentation of SIM materials effective as well as happy with how SIM has been implemented. However, normal classroom learning is still preferred over SIM learning in terms of increasing knowledge, increasing skills, imparting values and improving attitudes. Normal classroom learning is also preferred choice in comparison to SIM learning in terms of understanding English, Mathematics and Dzongkha subjects. In short, students, teachers, principals, district education officers, local government leaders and parents are happy with SIM programme as an Education in Emergency intervention but not as a better substitute for normal classroom learning during normal times.