Effect of Personalized, Computer-Based Instruction on Students’ Achievement in Solving Two-Step Word Problems

Adeneye O. A. Awofala
Department of Science and Technology Education
University of Lagos, Nigeria.

Abstract—This study investigated the effect of personalized, computer-based instruction using individual student interests and preferences on students’ achievement in solving two-step word problems. 80 junior secondary school year three students were grouped by ability level based on pre-test scores, then randomly assigned to a personalized or non-personalized version of the computer-based instruction on two-step word problems. Students made significant pre-test-to-post-test gains across treatments and the personalized treatment yielded a significant achievement difference over the non-personalized one. Significant two-way interactions for treatment by ability level, treatment by test occasion and ability level by test occasion were recorded. Also, a significant three-way interaction reflected that personalized high-ability students, non-personalized high-ability students, and non-personalized low-ability students improved less from pre-test to post-test than personalized low-ability students.

Keywords—Personalized, computer-based instruction, achievement, two-step word problems

I. INTRODUCTION

Word problems are those problems in which mathematical concepts and principles are expressed in everyday plain language, as different from purely formal mathematical symbols, signs, terminologies and expressions. Students often dread word problems and consider them distasteful and anxiety-inducing tasks in the mathematics classroom [30]. In Nigeria, evidence abounds that students perform poorly on solving routine word problems [22], [2]. A major cause of the under achievement appears to stem less from a lack of computational skills than from the inability to understand the problem and translate it into a mathematical expression [2], [17]. Besides, limited experience with word problems [5], lack of motivation to solve word problems [15], and irrelevance of word problems to students’ lives [13] contribute in no small measure to this poor performance.

Although word problems either routine or non-routine are difficult for students, they play a dominant role in mathematics education, and it is essential to seek improved ways of making the context of the problem more meaningful to students’ real-life situations. Research indicates that the closer the problem context is to students’ real-life situations, the more likely they will be able to comprehend and solve the problem [17]. Thus, linking classroom mathematics word problem-solving to students’ real life experiences may be to provide rich, meaningful contexts that situate both problems and the associated mathematics operations in familiar contexts. The result can lead to enhanced student thinking and make instruction in mathematical word problems more personally relevant to the students.

Personalizing mathematics word problems involves incorporating selected information with students’ personal preferences and interests into the problem context [6], [29], [17, 19]. Many studies on personalization assessed student learning outcomes using paper-based personalized mathematics instruction [6], [19]. In fact paper-based group personalization (tailoring problems to whole-class rather than individual content) as opposed to paper-based individual personalization was preferred because it was easier to construct and needed no computer system for its implementation.

However, one major limitation of paper-based personalization is that it is time-consuming to develop and implement individual personalized mathematics problems on paper. As a computer-based instructional practice, personalization has continued to need additional research in our increasingly technological world. A possible explanation for this phenomenon may be limited access to computers in schools to enhance its practical implementation. Adapting computer-based personalized instruction would reduce time investment and labour constraint in personalizing group worksheets and tests as well as promotes individualized personalization.

In a study by [9], fourth and fifth grade students worked with educational computer activities designed to teach arithmetic and problem-solving skills. Results indicated that personalization of the learning context produced increases in students’ intrinsic motivation and their depth of engagement in learning. In study by [3], fifth and sixth grade students scored significantly higher on mathematics word problems after receiving personalized computer-assisted lessons. They did better than peers without personalized instruction in solving standard problems and transfer problems, in recognizing rule procedures, and in task attitudes. Results of the study indicated that personalized contexts increase task motivation by describing applications of high interest to
learners and increase comprehension by helping learners interpret and inter-relate important information in the problem statements.

Reference [12] found that high school students preferred reading personalized stories to non-personalized ones, and that lower-ability group reported a significantly higher overall preference than higher ability group for the personalized stories. Reference [17] researched on the effects of personalization on 72 fifth grade Taiwanese students. The results of their study revealed that students made significant pre-test-to-post-test gains across treatments and scored significantly higher on personalized than on non-personalized post-test problems. Also, significant two-way interactions reflected greater pre-test-to-post-test improvement for lower-ability than for higher-ability students and a greater difference between scores on personalized over non-personalized post-test problems for lower-ability students.

In spite of the many successes of personalization whether paper-based or computer-based in promoting students’ performance on mathematics word problems, some investigations into its use have returned no positive results. Reference [6] found no significant increase in student achievement when paper-based personalization treatment was used despite student’s excitement on the personalized problems.

Reference [18] investigated the effects of personalized computer-based instruction in mathematics learning. The researchers found that although students made significant post test gains across treatments and scored significantly higher on arithmetic than on two-step word problems on the post test, the personalized treatment did not yield a significant achievement difference over the non-personalized one. Reference [8] found out that there were no significant differences between learners through personalized or non-personalised materials.

As noted, [17] assessed student performance using paper-based personalized mathematics instruction. Group personalization (tailoring problems to whole-class rather than individual interests) was implemented because it was easier to construct and because there was limited access to computers in the school in Taiwan to enhance individualized computer-based personalization. While study has shown the efficacy of paper-based personalization in promoting students’ achievement and self-efficacy in mathematics word problems in Nigeria [1], the efficacy of computer-based personalization on achievement in word problems in Nigeria is yet to be examined.

II. THE PROBLEM

The purpose of this study was to investigate whether individually personalized computer-based instruction on two-step word problems would improve student achievement. Specifically, the study examined the effect of two levels of personalized computer-based instruction (personalized, non-personalized) on the achievement of junior secondary school year three students on two-step word problems. Ability level and test occasion were considered as variables of interest in this study because of differential findings by ability and test occasion in previous research on personalization [18], [17], [20]. Ability level was considered at two levels (high, low) and test occasion at two levels (pre-test, post-test).

III. METHODOLOGY

1) Research design: This study adopted a pre-test – post-test equivalent control group experimental design, where R represents randomization of the participants, X represents exposure of a group to an experimental variable, C represents exposure of a group to the control or placebo condition and O represents observation or test administered.

Experimental group (R) O₁ X O₂ X gain = O₂ – O₁ O₁O₂ = pre-tests
Control group (R) O₀ C O₄ Cgain = O₄ – O₀ O₂O₄ = post-tests

The advantage of this design is that it controls the major threats to internal validity [16]. Student ability was examined as a moderator variable.

2) Sample and sampling procedure: Using purposive sampling technique, one private secondary school each was selected from the four geographical locations in Akinyele Local Government Area of Oyo State, Nigeria to make a total of four schools. Simple random sampling technique was used to pick one Junior Secondary School year three (JSSIII) class in each of the participating schools. Twenty mathematics students were randomly selected from each of the schools to make a total of 80 students. Their age range was 12-15 years with a mean age of 13.27 years and standard deviation of 2.5.

3) Research Instruments: Three instruments prepared, validated and used for the study are:

(1) Two-step Word Problem Achievement Test (TWPAT)
(2) Students’ Favourites Survey (SFS)
(3) Instructional Programme on Two-Step Word Problems (IPTWP)

1) Two-step Word Problem Achievement Test: The TWPAT had 15 items based on two-step word problem type of two-step problem. The TWPAT was made up of items of discrimination power of more than 0.40 and difficulty index of 0.40-0.60. It had a KR-21 reliability coefficient of .84. To determine the TWPAT score, only the results of the computations needed to solve the problems were scored. Each result on the TWPAT was scored as correct or incorrect only. One point was given to correct result for each step of the 15 two-step word problems, for a possible score range of 0 to 30 points. Examples of the test items are given below.

Example 1.
Tayo is reading a 545 page book. He has already 257 pages. If he reads 16 pages a day, how long will it him to finish the book?

Example 2.
Ade bought 15 oranges from the market at the rate of N5 per orange. If 5 oranges were spoiled, what is the total price of the remaining oranges?

2) Students’ Favourites Survey: A total of 20-items student favourites survey was used to determine the personal backgrounds and interest of the participants. Items included the name of student’s favourite places, friends, activities, sports, foods, and so forth. Students typed in one favourite response for each survey item. The SFS was designed based on literature.

3) Instructional Programme on Two-step Word Problems: The computer-based instructional programme on two-step word problems for this study was designed and developed with the assistance of a programmer using Macromedia Authorware. Two parallel versions of a computer-based instructional programme were designed and developed for solving two-step word problems involving the basic arithmetic operations in a mix. Each version required the same computational skills and used identical numbers but the problem context differed. The non-personalized version included standard two-step problem type from the students’ mathematics textbooks. A four-part strategy (understanding the problem, devising a plan, carrying out the plan, and looking back) for solving word problems based on the work of [23] was incorporated into the instructional programme on two-step word problems for both versions of the computer-based instructional programme.

Each student choices from the favourites survey were subsequently used by the software to convert the non-personalized version of the instructional programme into the personalized version. Instruction on the strategy for solving the two-step word problems also contained the four-part strategy and its application with appropriate worked examples and practice problems were provided. Six practice problems were given and the computer also provided instant feedback to students’ answers at each step, informing them whether their answers were correct or incorrect. If the student failed to correctly solve the problem, the computer would instantly provide the correct final answers on the screen and direct them to move on to the next question.

IV. PROEDURE

The experimental part of the study took place over two 45-minutes class periods on consecutive days, two weeks after the administration of the TWPAT as pretest. After the pretest had been scored, the students were grouped within each class by their pretest scores into higher-ability and lower-ability groups, and were randomly assigned within groups to either the personalized or the non-personalized versions of the computer-based instructional programme on two-step word problems. This resulted in 40 participants in the personalized treatment and 40 in the non-personalized treatment, with 20 high-ability and 20 low-ability participants in each group of 40.

On the first day of the experiment, all students filled out the Favourites Survey at the beginning of the computer-based instructional programme. For the personalized group, the software converted the non-personalized problems into the personalized content for the instructional programme using each response that students typed into the Favourites Survey.

The content of the instructional programme for the non-personalized group still remained the same despite their involvement in the filling of the Favourites Survey. All participants completed the instructional programme on two-step word problems that consisted of eight examples and six practices involving two-step word problems. On the final day, participants took the TWPAT as post-test. All summary sheets were printed and collected at the end of each class period by the computer laboratory teacher.

V. DATA ANALYSIS

The data collected were collated and analyzed using percentage, mean, standard deviation and 2x2x2 Analysis of Variance (ANOVA). Tukey HSD analysis was used in post hoc contrast. An alpha level of .05 was used for all statistical tests.

VI. RESULTS

Null hypothesis 1
There is no significant main effect of (i) treatment, (ii) ability level and (iii) test occasion on student achievement in two-step word problems.

| TABLE I |

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Test Occasion</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>High Ability</td>
<td>Low Ability</td>
</tr>
<tr>
<td>Personalized</td>
<td>M</td>
<td>9.98</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Non-Personalized</td>
<td>M</td>
<td>9.98</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3.2</td>
<td>1.3</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
<td>9.98</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>3.2</td>
<td>1.35</td>
</tr>
</tbody>
</table>

Mean number correct by variable:

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Ability</th>
<th>Test Occasion</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personalized</td>
<td>11.82</td>
<td>High</td>
<td>11.61</td>
<td>Pre-test</td>
</tr>
<tr>
<td>Non-personalized</td>
<td>6.32</td>
<td>Low</td>
<td>6.40</td>
<td>Post-test</td>
</tr>
</tbody>
</table>
The interaction for treatment by test occasion revealed that students made significant improvement from pre-test to post-test on personalized treatment (M = 6.18 Vs 11.82) than non-personalized treatment (M = 6.178 Vs 6.34). This improvement from pre-test to post-test on personalized treatment was strong enough to produce this significance (p<.001) as indicated by the test of simple effect using post hoc Tukey HSD analysis.

The interaction for treatment by ability level revealed that while high ability students averaged 9.98 each on the pre-test for both personalized and non-personalized treatment, the high ability students who received personalized instruction (M = 13.23) outperformed those who received non-personalized instruction (M = 10.12) on the post-test.

More so, despite comparable means (M = 2.38, M = 2.37) of low ability students on the pre-test for both personalized and non-personalized groups, the low ability students on the personalized treatment (M = 10.41) performed significantly better on the post-test than the low ability student on the non-personalized treatment (M = 2.53). This significant (p<.001) is confirmed by the test of single effect using post hoc Tukey HSD analysis.

The 2×2×2 ANOVA also yielded a significant threeway interaction for treatment by ability level by test occasion, F (1,72) = 8.16, MSE = 24.23, p<.05. This interaction indicated that the personalized low-ability students improved better from pre-test-to-post-test than personalized high-ability students, non-personalized high-ability students, and non-personalized low-ability students.

Personalized high-ability students had mean scores of 9.98 (66.5%) on the pre-test and 13.23 (88.2%) on the post-test, an improvement of 3.25 items correct. Non-personalized high-ability students had mean scores of 9.98 (66.53%) on the pre-test and 10.12 (67.46%) on the post-test, an improvement of 0.14 items correct. Non-personalized low-ability student had mean scores of 2.37 (15.8%) on the pre-test and 2.53 (16.87%) on the post-test, an improvement of 0.16 items correct, whereas personalized low-ability students had mean scores of 2.38 (15.87%) on the pre-test and 10.41 (69.4%) on the post-test, an improvement of 8.03 items correct.

A post hoc Tukey HSD analysis showed that the pre-test-to-post-test improvement for high-ability students between personalized and non-personalized treatment was not statistically significant (p = .30), whereas the pre-test-to-post-test improvement for low-ability students between personalized and non-personalized treatment was significant (p<.05).

### DISCUSSION

The results of this study indicated significant main-effect of treatment, ability level and test occasion. Significant two-way interactions were obtained for treatment by ability level; treatment by test occasion and ability level by test occasion. This latter interaction reflected higher gains by low-ability students than by high-ability students.

### Table II

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>Df</th>
<th>MSE</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment (A)</td>
<td>1434.63</td>
<td>1</td>
<td>1434.63</td>
<td>84.39</td>
<td>.000</td>
</tr>
<tr>
<td>Ability Level (B)</td>
<td>1385.12</td>
<td>1</td>
<td>1385.12</td>
<td>32.45</td>
<td>.001</td>
</tr>
<tr>
<td>Test Occasion (C)</td>
<td>285.72</td>
<td>1</td>
<td>285.72</td>
<td>42.14</td>
<td>.000</td>
</tr>
<tr>
<td>A×B</td>
<td>34.18</td>
<td>1</td>
<td>34.18</td>
<td>11.02</td>
<td>.001</td>
</tr>
<tr>
<td>A×C</td>
<td>32.45</td>
<td>1</td>
<td>32.45</td>
<td>10.34</td>
<td>.001</td>
</tr>
<tr>
<td>B×C</td>
<td>33.78</td>
<td>1</td>
<td>33.78</td>
<td>10.87</td>
<td>.000</td>
</tr>
<tr>
<td>A×B×C</td>
<td>24.23</td>
<td>1</td>
<td>24.23</td>
<td>8.16</td>
<td>.03</td>
</tr>
<tr>
<td>Error</td>
<td>1555.06</td>
<td>72</td>
<td>21.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4783.17</td>
<td>79</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sig. = Significant (p<.001)
The finding that the personalized treatment yielded a significant difference over the non-personalized treatment is consistent with the results in several personalization studies [1], [19], [20], [21], [11], [31], [3], [15], [10], [25] but at variance with those of (e.g. [8], [18], [6]). This research result, points to a similar conclusion in many previous studies on personalization.

First, the students’ greater familiarity with the personalized situations or content may have contributed to the treatment’s success. This familiarity may have reduced the students’ cognitive load in conceptualizing and processing the elements of the problem and may have enabled the students to solve the problem with less difficulty.

Second, the greater interest or motivation resulting from personalization may have been a factor in the treatment’s success. Students showed increased motivation when they saw their names or favourite things included in a problem. This was evident in the comments made by them while studying the instructional programme on two-step word problem. Comments such as “Hey, this includes my name” or “These problems are interesting” and the smiles that followed were taken as signs of increased student interest. This may have energized student to persevere on solving the problems.

Third, the relatively old age of the students may have contributed to the positive results of the present study. Studies have shown that older children in elementary school benefited greatly from personalisation of mathematics word problem than younger children [6], [11]. This is attributed to the fact that older children possess more developed schemata for processing information in a real-world context [4]. Age may be a determining factor in the choice of technique(s) to stimulate student interest in mathematics problem solving. While higher grade levels are noted for increasingly difficult Mathematics problems, the complexity of these problems may enhance personalisation strategy to influence student word problem achievement. Most studies that found positive effects for personalization (as indicated above) took place at upper elementary or middle grades. The present study dealt with Junior Secondary School year three (- an equivalent of ninth-grade) students and found relationship between personalization and student scores.

The significant differences for test occasion supported the claim that computer-based instruction can increase low-ability students’ mathematics achievement [18], [14], [29], [7].

Some research data suggest caution on overdependence on familiar problem contexts in mathematics instruction and assessment. Too much interest in problems may not only be detrimental to some students, who may incorrectly assume that they have attained correct answers [25], but distract some students, particularly girls [24], [30] and reduce transfer of learning to less familiar problem settings [18]. In line with [18], [9], the present study seems not to support these concerns, (at least, the latter) especially for low-ability students.

Prior to treatment, low-ability students assigned to the personalized treatment and those assigned to the non-personalized treatment had similar scores (15.87% and 15.8% respectively) on the non-personalized two-step word problems on the TWPAT. Following treatment, low-ability students on the personalized instructional programme scored significantly higher (69.4% to 16.87%) on the TWPAT than those on the non-personalized instructional programme. This improvement for low-ability student in the personalized treatment on non-personalized TWPAT items may be an indication for greater transfer of learning for personalized treatment engaged in personalized items than for non-personalized treatment engaged in non-personalized items during computer-based instruction.

VIII. CONCLUSIONS

The present study has implications for educational practice in Nigeria and elsewhere. Personalized computer-based instruction is effective in increasing students’ performance to solve two-step word problems. It is a catalyst for low-ability student performance to solve arithmetic problems. One application of the achievement results is for mathematics teachers in Nigeria and elsewhere to learn the interests and preferences of their students and incorporate these interests into their mathematics problems and instruction. It is also important that the content of new mathematics textbooks in Nigeria and elsewhere is made appealing to students by carefully attending to personalize and interesting problem context. Teachers of mathematics should learn to incorporate computer-based instruction into their teaching to support students’ learning and facilitate their performance.

However, one major limitation of this study was the smallness of the sample size. The sample size was reduced due to the few numbers of computers available at the schools during the period of the study. Twenty computers each were available in each of the four schools that participated in the study. In fact, availability of computer was a major criterion used in the selection of schools that took part in the study. Although, this greatly reduced the sample size, it did not limit the power of the data used to determine the effectiveness of the treatment.

REFERENCES


