The Effects of Adaptive Feedback on Student's Learning Gains

Andrew Thomas Bimba[1], Norisma Idris[2]
Department of Artificial Intelligence,
University of Malaya
Kuala Lumpur, Malaysia

Ahmed Al-Hunaiyyan[3]
Computer & Information Systems Department,
College of Business Studies,
Public Authority for Applied Education and Training (PAAET), Kuwait

Salwa Ungku Ibrahim[4], Naharudin Mustafa[5], Izlina Supa'at[6], Norazlin Zainal[7], Mohd Yahya Ahmad[8]
Center of Science Foundation Studies, University of Malaya,
Kuala Lumpur, Malaysia

Abstract—There is an increase in the implementation of adaptive feedback models, which focus on the relationship between adaptive feedback and learning gains. These literatures suggest that the complex relationship between feedback, task complexity, pedagogical principles and student's characteristics affect the significance of feedback effects. However, current studies have shown insufficient research on the effect of adaptive feedback characteristics on student's learning gains. Thus, there is a need to investigate the effect of multiple adaptive feedback characteristics on student's learning gains. The adaptive feedback model proposed supports the retrieval of appropriate feedback for students based on established weights between related concepts. In comparing three experimental groups, students who were provided with adaptive feedback showed learning gains and normalized learning gains of 0.87 and 0.05 over the normal feedback group, with 0.97 and 0.07 over the non-feedback group. This research yielded better outcomes than previous similar studies.

Keywords—*Authoring tools and methods; evaluation of CAL systems; intelligent tutoring systems; teaching/learning strategies; pedagogical issues*

I. INTRODUCTION

With progressions in educational innovations, e-learning systems have advanced to deliver learning conditions to the privileged and under-privileged individuals so they can study at their own speed. The accomplishment of these frameworks depends on engaging experience, timely and precise feeedback to the students [1] . One of the key element in learning support is to provide distinctive feedback based on an individual's requirement [2]. Customizing feedback according to student's qualities and other external parameters is a promising method for executing adaptation in computer-based learning environment [3], [4]. Adaptive feedback unlike generic feedback is dynamic, as students work through instructions, different students will receive diverse information [5]. Tending to this need, numerous researchers have proposed different ways to deal with assisting students in learning [6]. As a result, they have identified gaps and have been developing various frameworks and educational systems that are able to analyze student learning and provide adaptive feedback. These literatures suggests that the complex relationship between feedback, task complexity, pedagogical principles and student's characteristics affect the significance of feedback effects. The detailed categorization of previous methods used in providing multiple adaptive feedback to students, has been presented in our earlier work [7]. However, in this section, we briefly review recent adaptive feedback implementations which are similar to our previous work [7], with an addition of student's learning gains.

The author in [8], developed an intelligent tutoring system (ITS), using natural language processing and dialog to assist students in comprehending Newtonian Physics. The tool *Auto-Tutor*, provides feedback which consists of hints, prompts, and assertions are provided in form of dialogs according to a student's knowledge level. The main objective of the feedback provided by *Auto-Tutor*, is to simulate the construction of knowledge based on the constructivist principle [8]. Experimental results, show a learning gain of 0.8 between a control and experimental group.

On the other hand, *Guru* is an ITS which supports conversation with students while solving exercises in high school biology [9]. Similar to *Auto-Tutor*, *Guru* provides feedback in form of a dialog based on the student's knowledge level. However, *Guru* does not provide feedback based on any pedagogical principle or learning theory [7]. While comparing the control and experimental groups in an experiment, the effect of this feedback strategy on student's learning gains was 0.72 .

Gerdes's tutor, is developed as an interactive functional programming tutor for Haskell programming language [10]. The feedback provided by *Gerdes's tutor*, does not involve any characteristics of the students, however, the hint provided is generated automatically from a syntax tree of the model solution [7]. The hints provided are in form of steps, from a less reviling hint to a more detailed description of the solution.

A multimedia based ITS known as *Wayang Outpost*, is aimed at assisting student's in solving mathematical problems [11]. It provides feedback in form of hints, which are based on the learner's cognitive profile. These hints are provided based on the modality, contiguity, and animation multimedia learning principles. The cognitive apprenticeship principle is also utilized with the aim of encouraging students to accomplish more difficult task. However, an experiment to determine the

effect of this feedback strategy did not yield any significant results [7].

DeepTutor is another dialog-based intelligent tutoring system that uses scaffolding to improve student's knowledge during problem-solving [12]. *DeepTutor* utilizes the students' knowledge level in order to determine the type and frequency of feedback [12]. *DeepTutor* provides scaffolding and a sequence of progressive hints, based on the student's knowledge level as articulated in the student model [13], [12]. Experiments show a learning gain of 0.79 for students in the experimental group as compared to the control group.

The author in [14], developed a framework for generating a generic prompting principle and prompts with the ability to support learning and skill acquisition for novices solving ill-structured problems. The feedback provided by the framework is based on the different levels of problem difficulty. It utilizes meta-cognitive principles and scaffolding techniques to support learning in ill-structured problem-solving context [14]. However, this framework has not been evaluated to determine the effect of the proposed feedback strategy on student's learning gains.

The Paired Associate Deterministic Learning Task (PADL) was designed to determine if feedback information which is limited to positive and negative, is suitable in a deterministic learning process [15]. The feedback in PADL is provided based on the student's solution with motivational factor in the form of an expected monetary reward. Experiments indicate that the learning process is influenced by the type of feedback provided and the expected monetary reward. But, the feedback technique used was not evaluated for its effect on student's learning gains.

Computational Thinking using Simulation Modeling *CT-SiM*, is a learning tool that assist in learning science and computational thinking in middle school [16]. The feedback provided by *CTSiM*, is based on the student's proficiency which is evaluated based on the measure of the student's effectiveness and coherence. The hints are provided based on the scaffolding strategy using a conceptualized conversational dialog [16]. Experiments indicates positive learning gains of 0.29 between the experimental and control groups.

In an attempt to assist students in problem-solving, Negotiation-based adaptive learning system *NALS* was implemented [17]. *NALS*, uses student's solution to provide 3 level of hints and worked-out examples. The self-regulated learning principle is used for providing feedback, in order foster independence and lead students to seek help only when it is needed. Experiments indicate that *NALS*, helps in regulating student's help-seeking behaviors [17]. However, there was no study on the effect of this adaptive feedback technique on student's learning gains.

ALICE:fractions is an interactive mathematics textbook for helping students understand fractions [18]. Feedback in this tool is provided based on the student's answers. The principle of self-regulated learning is used in providing task level feedback in *ALICE:fractions*. Mostly, in *ALICE:fractions*, feedback is provided in form of text or visualization. Similar to other implementations, *ALICE:fractions* did not evaluate the effect of the feedback provided on student's learning gains.

The author in [1] developed a technique which provides feedback to students based on their performance, collaborative learning behavior, engagement, and understanding of concepts. This technique was reported to significantly improve student's performance. However, the student's learning gains were not considered in evaluating the system developed. Additionally, the complex relationship between task complexity, pedagogical principle, and student characteristics was not taken into consideration.

Paraprofessionals focus on educating students with severe disabilities. The author in [19] focused on the techniques that make paraprofessionals more effective in providing feedback. In their study the demonstrated that delayed, video-based performance feedback is an effective and efficient means of providing feedback. Based on the student's performance, simultaneous prompting and least-to-most (LTM) prompting are provided as feedback.

The Life-Saving Instruction for Emergencies (LIFE) app was developed by [20] to provide health care providers with self-regulatory learning content which is independent of classroom tutoring. A simulated medical emergency is made available to health care providers using a smartphone-based game, through which their management skills is been assessed. An immediate feedback is provided at every wrong attempt. The effectiveness of the feedback was measured based on the learning gains of an experiment and controlled group. The experimental group which were provided with adaptive immediate feedback showed a learning gain of 0.09.

The author in [21] developed a learning analytic dashboard that provides visualized feedback and adaptive support for face-to-face collaborative argumentation (FCA). Students utilize the dashboard to monitor their FCA process and the instructors provide real-time immediate feedback based on the student's performance.

Table I shows a summary of adaptive feedback implementations, indicating the feedback characteristics and learning gains. Based on these current studies, there are insufficient researches on the positive effect of multiple adaptive feedback characteristics on student's learning gains. A reason for this could be the complex nature of representing multiple adaptive feedback characteristics. Thus, these research aims at investigating the effect of multiple adaptive feedback characteristics on student's learning gains. To achieve the aim of this research an adaptive feedback model and algorithm is proposed in Section 2. In Section 3, the evaluation methods and processes are discussed. While the results of the evaluation process is presented in Section 4, and discussed in Section 5. Finally, the findings, limitations of the research, and future works are presented in Section 6.

II. PROPOSED ADAPTIVE FEEDBACK MODEL

Information in a computer-based learning environment can be regarded as models. The three most imperative models are the pedagogical model, domain model, and the student model [7]. The pedagogical model represents information and ways of teaching. The particular knowledge represented in the pedagogical model is based on pedagogical principles or learning theories. This principle or theory decides the successful teaching methods, instructional methods, sequence of activities, feedback types, and assessments modeled by the

TABLE I. ADAPTIVE FEEDBACK IMPLEMENTATIONS

S/N	Feedback Implementation	Feedback Characteristics	Learning gains	Normalize learning gains
1	AutoTutor (2012) [8]	Student's knowledge level, hints, student's solution, prompts, constructivist principle, Sequence of feedback	0.8	N/A
2	Guru (2012) [9]	Student's knowledge level, Scaffolding and sequential hints (domain knowledge), Based on a constructivist scaffolding strategy, Levels of details, Active learner.	0.72	
3	Gerdes's,Tutor (2012) [10]	Local to global	N/A	N/A
4	Wayang Outpost (2014) [11]	Cognitive apprenticeship, Step-by-step hints	Not significant	N/A
5	DeepTutor (2015) [12]	Student's knowledge level, scaffolding and sequential hints, based on a constructivist scaffolding strategy, levels of details and active learner.	0.79	N/A
6	Smy (2016) [14]	prompts, meta-cognition, scaffolding, and problem's level of difficulty.	N/A	N/A
7	PADL (2017) [15]	Monetary reward, positive and negative feedback, and learner's solution	N/A	N/A
8	CTSiM (2017) [16]	Student's proficiency, incorrect solutions, scaffolding, and hints	0.29	N/A
9	NALS (2018) [17]	Student's solution, hints and prompts, worked example, levels of detail, and self-regulated learning.	N/A	N/A
10	ALICE:fractions (2018) [18]	Student's answer and self-regulated learning.	N/A	N/A
11	[1]	Performance, collaborative learning behavior, engagement, and understanding of concepts.	N/A	N/A
12	[19]	Performance, simultaneous prompting, and least-to-most (LTM) prompting.	N/A	N/A
13	LIFE [20]	Performance, immediate feedback, and incorrect attempt.	0.09	N/A
14	[21]	Immediate feedback and student's activities.	N/A	N/A

learning environment. In a different circumstances, the domain model, is an aspect of the pedagogical model, which represents the knowledge of the subject been learned. The domain model represents concepts, learning materials, facts, problems, solutions, feedback, rules equations, etc. Finally, the student model shows information about the student's knowledge of the domain, learning style, interactions with the system, response to feedback, emotional state, performance etc. These data decides the characteristics of the students while interacting with the domain and pedagogy.

In this research, we center on the cognitive apprenticeship pedagogical principle and the physics domain. The challenge, is to show the properties of the cognitive apprenticeship principle that influence the type, timing, goal, and sequence of feedback, relating it to the characteristics of the domain and student models. Fig. 1, shows the Pedagogy, Domain, and Learner (PDL) model which were introduced in our earlier work [22]. The representation of the characteristics of all 3 models relating to adaptive feedback is shown in Fig. 1. The knowledge shown in the proposed PDL model is based on the OAR model. The nodes of the concept network in Fig. 1, represents concepts in the pedagogy, domain and student model in a computer-based learning environment. The arcs represent the relationship between concepts and their traits.

A. Adaptive Feedback Algorithm

The purpose of the adaptive feedback algorithm is to determine the appropriate feedback for different students based on adaptive characteristics [23]. The input to the adaptive feedback algorithm is the partial problem solving state PS_s. PS_s consists of attributes from the domain and student model. As show in Eq. 1, the problem solving state involves the student's cognitive style, student's knowledge level,goal (expected performance and completion time) problem's level of difficulty, and domain topic.

$$PS_s = (S_{cs}, S_{kl}, P_s, P_t, C_p, D_c) \tag{1}$$

The similarity between the current problem solving state PS_s and existing problem solving states $PS_s i$ in the knowledge base is determined iteratively. All the states that are above a certain threshold are stored as a list. The problem solving state with the highest cumulative weight is selected and its influence on the student's performance which has been defined by the knowledge bonding algorithm (Algorithm 1) is the output.

Thus, during a learning season when the problem solving state is determined, an appropriate feedback is provided based on existing knowledge of the most similar state in the knowledge base. This approach utilizes the content-addressed
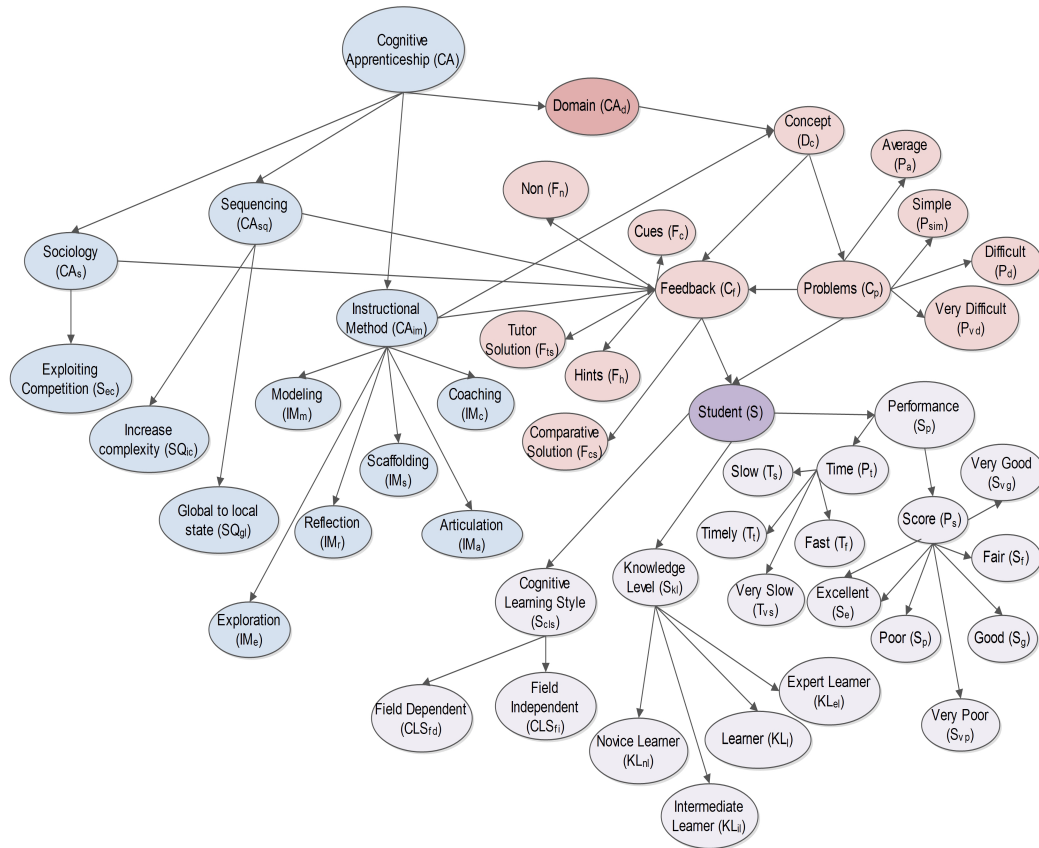
Fig. 1. PDL Model for Adaptive Feedback in Physics [22].

mechanism of a CKB for knowledge retrieval and manipulation which is enabled by the structural models.

III. MATERIALS AND METHOD

To establish the effectiveness of the PDL model in providing adaptive feedback, a group of pre-university (foundation studies) students are studied. The analysis is focused on the learning gains received by students receiving the PDL-based feedback as compared to traditional feedback and no feedback at all. The traditional feedback provides only feedback to the students without considering the adaptive characteristics of feedback.

A. Participants

The experimental process involves students distributed into three groups. The three groups consist of the traditional feedback group (Normal), no feedback (Non) group, and the experimental group (Adaptive). All students are allowed to take their regular classes. A pre-test before exposing students to the adaptive feedback tool is provided, followed by a problem-solving session with the adaptive feedback environment for four weeks, with each session lasting 30 minutes. After the experimental process, the students are provided with a post-test.

The pre-test and post-test comprise of 10 physics problems each in introductory mechanics with similar sub-topics and difficulty levels. While the problems provided by the PDL-based adaptive feedback system comprise of 160 problems. The students are required to solve 10 problems daily for the 4 weeks (5 days/week) period of the experiment. All the problems are provided by the physics experts.

The participants are 31 pre-university (foundation studies) students from Center of Foundation Studies of the University of Malaya taking physics course and are currently learning a topic on introductory mechanics. They are randomly categorised into three groups:

- The traditional feedback group (Normal) consist of students who solve problems with the adaptive feedback system, but are provided with the same type of feedback without considering the adaptive characteristics of feedback.

- The no feedback group (Non) involves students who undergo the problem-solving process with the adaptive feedback system with no feedback provided.

Algorithm 1: Adaptive Feedback Algorithm

Data: PS_s,time stamp
Result: R^{in}
read data;
create and initialize problem solving state list PS_{sL}
while PS_{si} *is available* **do**

    compute similarity $PS_s \sim PS_{si} = \left| \frac{A \cap A_i}{A \cup A_i} \right| =$

$$\begin{cases} 1 & PS_s = PS_{si} \\ (0,1) & PS_s \leftrightarrow PS_{si} \vee PS_s < PS_{si} \vee PS_s > PS_{si} \\ 0 & PS_s \neq PS_{si} \end{cases}$$

    ;
    if $PS_s \sim PS_{si} > TH$;   /* TH is the
    similarity threshold */
    then
        retrieve PS_{si};
        add to problem solving state list
        $PS_{sL} = PS_{sL} + PS_{si}$;
        current section becomes this one;
    else
        Create random $PS_s i$

determine highest PS_{si};
output the influence of related concepts of the highest
PS_{si}, $R_i^{in} \subseteq \bigcup_{i=1}^{|R_i|} R_i^c$;

TABLE II. Participant's Demographics

S/N	Size	Total
Age		31
18 Years	31	
Gender		
Male	11	31
Female	20	
Others	0	
Major		
Physical Science	11	31
Life Science 1	11	
Life Science 2	9	

- The experimental group (Adaptive) comprises of students who undertake the problem-solving process while they are provided with adaptive feedback based on the PDL-model.

The 31 participants are all 18 years old with 64.52% female and 35.48% male as shown in Table II. The participants where taken from three majors of Physical Science, Life Science 1, and Life Science 2. The three classes of students are then assigned to the Adaptive, Normal, and Control groups as shown in Table IV

All participants are enrolled in the pre-university (foundation studies) program at the same time and are attending the same classes in introductory mechanics. The problems provided are from four main topics of Static, Linear Motion, Rotational Motion, and Gravitation in introductory physics.

The problems are based on three levels of difficulty which includes easy, moderate, and difficult. The distribution of the problems based on the topics and difficulty level is shown in Table III.

B. Evaluation Method

Cross-disciplinary researches that involves artificial intelligence in education can be difficult to conduct due to the restrictions placed on research with human subjects. Experimental controls are restricted, and arbitrary sampling is difficult to achieve in a classic classroom setting. As computer scientists, we ought to have the expertise to construct valid, compelling studies of student's learning under the difficult conditions imposed by researches based on human subjects [24].

1) Learning Gains: Student's learning with a computer-based learning approach is usually evaluated using the relationship between a pre-test and a post-test scores to determine the learning gains and normalized learning gains [25]. The tests given prior to the provision of computer-based instructions or interventions are pre-tests, while the tests administered after the intervention are post-test [26]. In general, students who tend to perform well during a pre-test, do well in the post-test. The learning and normalized learning gains utilizes pre-test and post-test to evenly compare students' learning among experimental and control groups [27]. Equations 2 and 3 indicate that both the learning gains and normalized learning gains incorporate pre-tests (pre) and post-tests (post).

The normalized learning gains was introduced because, the learning gains provided an unfair advantage to groups with low pre-test scores. Most assessments involve problems with different levels of difficulty, and some information are easier to assimilate than others. Thus, a student with a low pre-test score has a wider scope of difficulty levels to learn than a student with high pre-test scores. The normalized learning gain was developed to rectify this uneven advantage [28].

2) Analysis of Covariance: In statistics and probability, there is a possibility that a trend is seen in several groups of data but later disappear when the data is combined, this phenomenon is called the Simpson's paradox. The paradox shows that a statistically significant relationship in a certain direction can result to an opposite direction when investigated with a covariant [29]. Since most investigations do not arbitrary assign students to experimental groups and the pre-test scores are usually related to the post-test scores, learning gains can show significant difference which favors a certain group but another group may be more effective when using the pre-test as a covariate. Analysis of Covariance (ANCOVA) eliminates the impact of a covariate to establish a fair comparison among different groups.

ANCOVA assesses whether the means of a dependent variable are equivalent throughout the dimensions of an independent variable, while statistically managing the impacts of continuous variables that are not essential, known as covariates [30]. As shown in Equation 4, an ANCOVA model expects a linear relationship between the dependent variable and the covariate.

$$y_{ij} = \mu + \tau_i + B(x_{ij} - \overline{x}) + \epsilon_{ij}. \tag{4}$$

TABLE III. DISTRIBUTION OF PROBLEMS IN INTRODUCTORY PHYSICS

S/N	Topic	Difficult	Intermediate	Easy	Total
1	Static	5	2	1	8
2	Linear Motion	26	32	26	84
3	Rotational Motion	10	12	16	38
4	Gravitation	10	10	10	30
5	Total	51	56	53	160

$$Learning\ gains\ lg = post - pre \tag{2}$$

$$Normalized\ learning\ gains\ nlg = \frac{learning\ gains}{possible\ learning\ gains} = \frac{(post - pre)}{(100\% - pre)} \tag{3}$$

TABLE IV. DISTRIBUTION OF STUDENTS AMONGST EXPERIMENTAL GROUPS

S/N	Major	Adaptive	Normal	Non	Total
1	Physical Science	3	5	3	11
2	Life Science 1	5	4	2	11
3	Life Science 2	4	3	2	9
4	Total	12	12	7	31

TABLE V. DISTRIBUTION OF STUDENT'S RESPONSE TO THE EXPERIMENT

S/N	Group	Pre-test	Adaptive Tool	Post-test	All
1	Adaptive	18	12	12	12
2	Normal	16	12	12	12
3	Non	16	7	9	7
4	Total	50	31	33	31

In Equation 4,

- dependent variable y_{ij} is the jth observation for the ith group.

- covariate x_{ij} is the jth observation of the covariate the ith group.

- grand mean is μ

- global mean of covariant is x

- the effect of the ith level of the independent variable is τ_i

- line slope is B

- associative non observed error for the jth observation for the ith group is ϵ_{ij}

In conducting an ANCOVA, there are some assumptions that are tested. First there is a need to determine that no outliers exists. Then, the observations in the dependent variable should be normally distributed for each level of the independent variable. This test is known as the assumption of normality. This assumption requires multiple conditional distributions of the dependent variable for every combination of the covariate to be normally distributed. If the population distributions are not normal and the sample sizes are small, the significant values obtained with ANCOVA may be invalid. Another assumption that is tested in ANCOVA, is the homogeneity of regression slopes. The regression lines for individual groups are assumed to be parallel, having the same slope. The failure to meet this assumptions implies that there is an interaction between the covariant and the independent variable. Usually this assumption is evaluated by the F-test, if the value is significant then the assumption of homogeneity of regression

slopes is violated. The final assumption considered in this experiment is the homogeneity of variance. This test shows the relationship between the dependent variable and the covariate for the entire dataset, ignoring the various groups of the independent variable. Thus, if there is a positive relationship between the covariate and the independent variable of one group, then there should be a possible relationship for the rest of the groups.

IV. ANALYSIS OF RESULTS

The pre-test and post-tests scores are rated on a 100-points scale, where 0 indicates a poor performance and 100 indicates an excellent performance. The rating is done by physics experts who provided the problems to be solved. The assessment focused on the workings towards a solution, identification of known, unknowns, equations and principles along with the correct answer. At the beginning of the experiment, 63 students were invited, 50 turned up for the pre-test and registered to use the adaptive learning environment. Then, out of the 50 students, 33 students turned up for the post-test. In the end, only 31 students did the pre-test, used the adaptive learning tool, and also participated in the post-test. Table V, show the detailed break-down of the students responses to the experiment according to the experimental groups.

1) Results of Student's Learning Gains: In order to determine the effect of proposed adaptive feedback model on student's learning gains, the results of the pre-test and the post-test of each group is collected as shown in Table VI. The pre-test and post-test scores are based on the average scores of the 10 problems solved by each student. The learning gains and normalized learning gains are calculated according to Equations 2 and 3. The average learning gains of each experimental group and a comparison between the average

TABLE VI. STUDENT'S PRE-TEST AND POST-TEST SCORES

S/N	Student ID	Major	Group	Pre-test Score	Post-test Score	Learning Gains	Normalized Learning Gains
1	0073	Life Science 2	Adaptive	0.00	9.90	9.90	0.10
2	0036	Physical Science	Adaptive	2.00	16.00	14.00	0.16
3	0057	Life Science 1	Adaptive	8.70	25.20	16.50	0.28
4	0041	Physical Science	Adaptive	2.00	20.40	18.40	0.21
5	0048	Life Science 1	Adaptive	5.50	28.00	22.50	0.30
6	0061	Life Science 2	Adaptive	1.00	13.70	12.70	0.14
7	0072	Life Science 2	Adaptive	3.00	24.50	21.50	0.25
8	0056	Life Science 1	Adaptive	6.50	32.40	25.90	0.35
9	0045	Life Science 1	Adaptive	7.80	20.90	13.10	0.23
10	0051	Life Science 1	Adaptive	0.70	17.70	17.00	0.18
11	0059	Life Science 2	Adaptive	4.00	21.50	17.50	0.22
12	0032	Physical Science	Adaptive	8.00	31.50	23.50	0.34
13	0037	Physical Science	Normal	3.00	10.00	7.00	0.10
14	0067	Life Science 2	Normal	1.00	7.60	6.60	0.08
15	0055	Life Science 1	Normal	10.40	10.00	-0.40	0.11
16	0074	Life Science 2	Normal	6.25	7.70	1.45	0.08
17	0031	Physical Science	Normal	9.00	17.70	8.70	0.19
18	0049	Life Science 1	Normal	12.50	10.50	-2.00	0.12
19	0054	Life Science 1	Normal	5.80	29.00	23.20	0.31
20	0047	Life Science 1	Normal	13.00	29.50	16.50	0.34
21	0044	Physical Science	Normal	3.50	9.50	6.00	0.10
22	0070	Life Science 2	Normal	0.00	2.40	2.40	0.02
23	0030	Physical Science	Normal	9.30	19.50	10.20	0.21
24	0043	Physical Science	Normal	6.00	11.90	5.90	0.13
25	0064	Life Science 2	Non	10.55	13.30	2.75	0.15
26	0046	Life Science 1	Non	5.50	13.5	8.00	0.14
27	0033	Physical Science	Non	6.50	10.00	3.50	0.11
28	0052	Life Science 1	Non	11.90	7.00	-4.90	0.08
29	0042	Physical Science	Non	6.30	11.80	5.50	0.13
30	0038	Physical Science	Non	6.50	20.40	13.90	0.22
31	0062	Life Science 2	Non	0.50	10.15	9.65	0.10

TABLE VII. COMPARISON OF STUDENT'S LEARNING GAINS

S/N	Group	Learning Gains	Normalized Learning Gains
1	Adaptive	16.24	0.21
2	Normal	7.98	0.16
3	Non	6.55	0.14
4	Adaptive Vs. Normal	8.72	0.05
5	Adaptive Vs. Non	9.69	0.07
6	Normal Vs. Non	1.43	0.02
7	Adaptive Vs. (Normal & Non)	8.98	0.06

learning gains are presented in Table VII. As expected, each experimental groups had a positive learning and normalized learning gains.

2) One-way ANCOVA Results: In this study, ANCOVA is used with a view to eliminate bias for a certain experimental group. The Statistical Package for the Social Science (SPSS)

tool was used to analyze the student's pre-test and post-test scores shown in Table VI. The ANCOVA allows us to see if there is a difference in the dependent variable (post-test) by the levels of the independent variable (group) while controlling for the effect of the covariate (pre-test). The independent variable used is the experimental group (group), has three levels which are adaptive, normal, and non. The dependent variable is the post-test, which is measured at a continuous level, while the pre-test is the covariate, which is also measured at a continuous level.

Table VIII, shows a descriptive statistics of the post-test according to the three experimental groups. It provides a summarized description of the post-test, for each group by presenting the mean, median, variance, range, standard deviation, skewness, and kurtosis. The mean post-test score of the adaptive feedback group was higher than the normal and non feedback groups.

In respect to ANCOVA models, data must be scrutinized to evaluate the existence of outliers, and affirm that it meets fundamental assumptions of normality, homogeneity of regression

TABLE VIII. POST-TEST DESCRIPTIVE STATISTICS

		Group		Statistic	Std. Error
PostTest	Adaptive	Mean		21.8083	1.99292
		95% Confidence Interval for Mean	Lower Bound	17.4219	
			Upper Bound	26.1947	
		5% Trimmed Mean		21.8815	
		Median		21.2000	
		Variance		47.661	
		Std. Deviation		6.90368	
		Minimum		9.90	
		Maximum		32.40	
		Range		22.50	
		Interquartile Range		10.88	
		Skewness		-.032	.637
		Kurtosis		-.640	1.232
	Non	Mean		12.3071	1.59091
		95% Confidence Interval for Mean	Lower Bound	8.4143	
			Upper Bound	16.2000	
		5% Trimmed Mean		12.1524	
		Median		11.8000	
		Variance		17.717	
		Std. Deviation		4.20916	
		Minimum		7.00	
		Maximum		20.40	
		Range		13.40	
		Interquartile Range		3.50	
		Skewness		1.129	.794
		Kurtosis		2.236	1.587
	Normal	Mean		13.7750	2.45058
		95% Confidence Interval for Mean	Lower Bound	8.3813	
			Upper Bound	19.1687	
		5% Trimmed Mean		13.5333	
		Median		10.2500	
		Variance		72.064	
		Std. Deviation		8.48904	
		Minimum		2.40	
		Maximum		29.50	
		Range		27.10	
		Interquartile Range		10.90	
		Skewness		.982	.637
		Kurtosis		.112	1.232

and variance [31]. Fig. 2, shows a box plot showing the overall pattern of the post-test scores for the experimental groups. The result from the box plot indicates that the adaptive and normal feedback groups have varying post-test scores, while the non feedback group had similar post-test scores. However, one outlier was identified in the non feedback group as seen in Fig. 2.

The results of the normality test is shown in Table X. The observations from the Shapiro-Wilk test shows that the post-test is normally distributed for the adaptive, normal, and non feedback groups. Furthermore, Fig. 3, 4, and 5, presents the Quantile-Quantile (Q-Q) plots for the three experimental groups. The ideal dataset should have all points on the trend line, however, the results presented from the Q-Q plots are not ideal.

To further explore the data, the homogeneity of regression slopes assumption is tested. It is assumed that the regression lines for the adaptive, normal, and non feedback groups are parallel. Table X, displays the results of the test for homogeneity of regression slopes using the general linear model in SPSS. The p-value (0.106) of the interaction term is non-significant (for $\alpha = 0.05$) indicating a non-violation of the
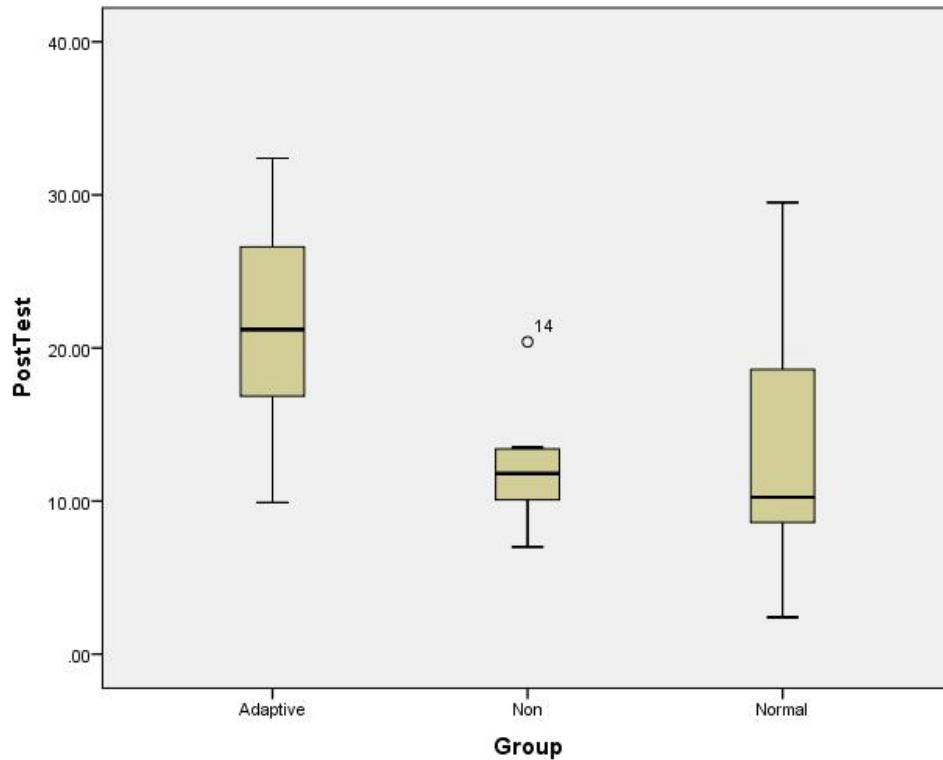
Fig. 2. Test for Outliers.

TABLE IX. TESTS OF NORMALITY

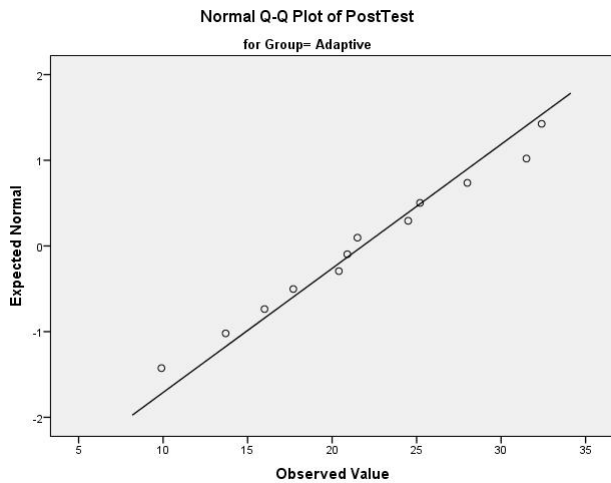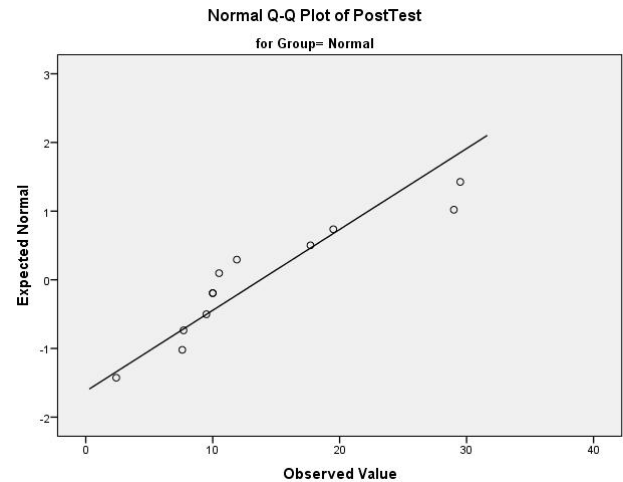|	Group	Shapiro-Wilk		
		Statistic	df	Sig.
PostTest	Adaptive	.978	12	.973
	Non	.914	7	.421
	Normal	.867	12	.060

Fig. 3. Adaptive Feedback Q-Q Plot for Post-Test.

Fig. 4. Normal Feedback Q-Q Plot for Post-Test.

assumption of homogeneity of regression slopes. Thus, there is no interaction between the pre-test and the experimental groups.

In this study, the last assumption tested was the homogeneity of variance. The Levene's test of homogeneity of variance was carried out. As shown in Table XI, this assumption is satisfied because $p(0.773) > \alpha(0.05)$

For testing the main hypothesis, a One-way ANCOVA was conducted to determine a statistically significant difference

TABLE X. TEST OF HOMOGENEITY OF REGRESSION SLOPE

Tests of Between-Subjects Effects
Dependent Variable: PostTest

| Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1111.712a	5	222.342	6.450	.001	.563
Intercept	952.967	1	952.967	27.645	.000	.525
Group * Pretest	169.187	2	84.594	2.454	.106	.164
Group	127.362	2	63.681	1.847	.179	.129
Pretest	268.347	1	268.347	7.785	.010	.237
Error	861.783	25	34.471			
Total	10467.783	31				
Corrected Total	1973.495	30				

R Squared = .563 (Adjusted R Squared = .476)
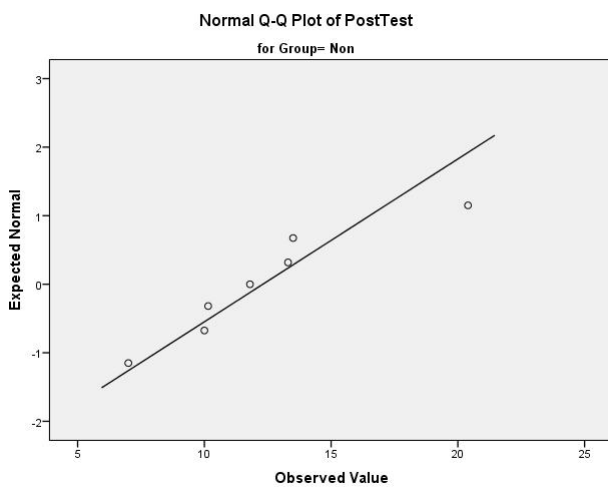
Fig. 5. No Feedback Q-Q Plot for Post-Test.

TABLE XI. TEST OF HOMOGENEITY OF VARIANCE

Levene's Test of Equality of Error Variances
Dependent Variable: PostTest

F	df1	df2	Sig.
.260	2	28	.773

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
Design: Intercept + Pretest + Group

between students provided with adaptive feedback, students provided with normal feedback and students with no feedback on a post-test controlling for pre-test. The results shown in Table XII, shows a significant effect ($p < 0.001$) of adaptive feedback on the post-test after controlling for the pre-test.

Considering the fact that a significant main effect was obtained and there are three levels of the independent variable, there is a need to conduct a post hoc to determine the significant pairwise differences. According to the results in Table XIII, there is a significant difference between adaptive and normal feedback groups ($p = 0.001$) and adaptive and non feedback groups ($p = 0.001$). However, there was no significant difference between the normal and non feedback groups ()$p = 1.000$).

V. DISCUSSION

The main aim of conducting this experiment is to determine the effect of adaptive feedback on student's learning gains. The three groups analyzed in this experiment are the adaptive, normal, and non feedback groups. To achieve a valid conclusion on the effect of adaptive feedback, two fundamental methods were used. First, the student's learning gains (LG) and normalized learning gains (NLG) were analyzed. Based on the results in Table VII, all the students in the three groups achieved positive learning gains and normalized learning gains after using the adaptive feedback tool. This is due to the interventions received by the students between the pre-test and post-test periods. However, despite the positive learning gains and normalized learning gains, the group which were provided with adaptive feedback ($LG = 16.24, NLG = 0.21$) had higher learning and normalized learning gains than those with the normal feedback ($LG = 7.98, NLG = 0.16$) and non feedback ($LG = 6.55, NLG = 0.14$). In comparing the three experimental groups, students who were provided with adaptive feedback showed learning gains and normalized learning gains of 8.72 and 0.05 over the normal feedback group, with 9.69 and 0.07 over the non feedback group. The results of this comparison shows that the student's who received adaptive feedback achieved superior learning gains than those who received normal and no feedback. In addition, student's who received normal feedback, had better learning gains than student who did not receive feedback. The results of this experiment implies that the adaptive feedback provided by the proposed PDL model has a positive effect on student's learning gains.

Previous researchers have conducted similar experiments to determine the effect of the feedback provided. Table XIV, shows the comparison of our results with the previous studies on a 10 point scale. The adaptive feedback provided by the proposed PDL model has more positive effect on student's learning gains than adaptive feedback models proposed in previous researches.

In view of eliminating any form of bias in our experimental results, a one-way ANCOVA was conducted to compare the effectiveness of the three feedback methods using the pre-test and post-test scores. The main hypothesis states that *the*

TABLE XII. ANALYSIS OF COVARIANCE

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	942.525a	3	314.175	8.228	.000	.478
Intercept	831.111	1	831.111	21.766	.000	.446
Pretest	392.304	1	392.304	10.274	.003	.276
Group	827.258	2	413.629	10.832	.000	.445
Error	1030.970	27	38.184			
Total	10467.783	31				
Corrected Total	1973.495	30				

R Squared = .478 (Adjusted R Squared = .420)

TABLE XIII. POST-HOC ANALYSIS OF COVARIANCE

Pairwise Comparisons
Dependent Variable: PostTest

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Adaptive	Non	12.232*	3.060	.001	4.422	20.043
	Normal	10.588*	2.646	.001	3.835	17.341
Non	Adaptive	-12.232*	3.060	.001	-20.043	-4.422
	Normal	-1.644	2.939	1.000	-9.147	5.859
Normal	Adaptive	-10.588*	2.646	.001	-17.341	-3.835
	Non	1.644	2.939	1.000	-5.859	9.147

Based on estimated marginal means
* The mean difference is significant at the .05 level.
b Adjustment for multiple comparisons: Bonferroni.

TABLE XIV. COMPARISON OF PROPOSED PDL MODEL AND PREVIOUS MODELS

S/N	Feedback Tool (Year)	Learning Gains	Normalized Learning Gains
1	AutoTutor (2012) [8]	0.8	N/A
2	Guru (2012) [9]	0.72	N/A
3	Wayang Outpost (2014) [11]	not significant	N/A
4	DeepTutor (2015)[12]	0.79	N/A
5	CTSiM (2017) [16]	0.29	N/A
6	LIFE (2020) [20]	0.09	N/A
7	Proposed PDL Model	0.90	0.06

adaptive feedback group will perform better than the normal and no feedback group in a post-test after using the adaptive feedback tool. During the analysis of the data, one outlier was identified in the non feedback group. The outlier was not removed from the dataset because there was no significant difference in the overall analysis with its exclusion. The Shapiro-Wilk test for normality showed that the post-test is evenly distributed between the groups with a non-significant value of $p(0.973) > \alpha(0.05)$ for adaptive, $p(0.60) > \alpha(0.05)$ for normal, and $p(0.421) > \alpha(0.050$ for non. The test for homogeneity of regression was conducted and the assumption was made with a non-significant value of $p(0.106) > \alpha(0.05)$. Then, the Levene's test for homogeneity of variance was carried out and a non-significant value of $p(.773) > \alpha(0.05)$ was obtained. This indicates that the null hypothesis of the error variance of the post-test is equal across all experimental groups. Since non of the assumptions tested were violated, the results obtained from the main ANCOVA is valid.

From the results, it is found that there was a significant difference in the mean post-test scores ($[F(2, 27) = 10.832, p <$

$0.001])$ between adaptive, normal, and non feedback groups. The partial Eta Squared value of 0.445 indicates the effect size. Based on the Cohen' s guideline the effect size is small. This implies that the variance in the post-test to a small effect, is explained by the experimental groups. The post hoc test indicated a significant difference between the adaptive and normal feedback groups ($p(0.001) < \alpha(0.05)$), adaptive and non feedback groups ($p(0.001) < \alpha(0.05)$). This indicates that the adaptive feedback technique provided by proposed the PDL model has a significant effect in improving student's post-test scores than the normal and no feedback techniques. Comparing the normal and feedback groups showed no significant difference ($p(1.000) > \alpha(0.05)$).

VI. CONCLUSION

The main objective of this research is *to evaluate the effect of the adaptive feedback model, on student's learning gains*. The effect of the adaptive feedback provided based on the proposed model, was evaluated to determine its effect on student's learning gains. Overall, the experimental findings

from calculating the students learning gains and a One-way ANCOVA, showed that the adaptive group who receive feedback based on the PDL model showed a statistically significant gain in learning outcomes as compared to the normal and non feedback groups. This suggests that adaptive feedback based on the PDL model can have a sustained positive impact on student's learning gains. Additional findings from the one-way ANCOVA showed that, there is no statistically significant difference in the post-test scores of students who are provided with normal feedback, as compared to those who had no feedback. However, the student's provided with normal feedback had better learning gains and normalized learning gains of 1.43 and 0.02, respectively, as compared to those without feedback. A comparison of the findings with other similar studies showed that the adaptive feedback provided by the PDL model is superior. As a conclusion, based on the aim of this case study, the provision of adaptive feedback using the PDL model is an effective tutoring strategy. Limitations of this study are moderate. First of all, a single case study design, of the effect of the adaptive feedback on students learning gains is not representative. In addition, more number of students might have been desirable, however, this was not possible due to the availability of the students. Besides the number of students, more interactions by the students with the adaptive learning tool is also desirable, but was not possible due to the nature of interaction with human subjects. This could be due to the difficulty in using the tool, as the student's are required to be familiar with writing equations in the latex format. Subsequently, from a science measurement perspective, while the adaptive feedback model fit reasonably well and indicated superior learning gains, additional characteristics of the pedagogy, domain, and student models need to be considered to determine the best fit. The proposed adaptive feedback model can be extended by adding additional characteristics of the student, domain and pedagogy. This new model can be evaluated to determine which characteristics has more effect to improve student's learning gains.

Acknowledgment

References

[1] M. A. Hassan, U. Habiba, H. Khalid, M. Shoaib, and S. Arshad, "An adaptive feedback system to improve student performance based on collaborative behavior," *IEEE Access*, vol. 7, pp. 107 171–107 178, 2019.

[2] E. G. Advisors, "Learning to adapt: A case for accelerating adaptive learning in higher education," *Education Growth Advisors*, 2013.

[3] A. A. Al-Hunaiyyan, A. T. Bimba, and S. Alsharhan, "A cognitive knowledge-based model for an academic adaptive e-advising system," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 15, pp. 247–263, 2020.

[4] S. Narciss, S. Sosnovsky, L. Schnaubert, E. Andres, A. Eichelmann, G. Goguadze, and E. Melis, "Exploring feedback and student characteristics relevant for personalizing feedback strategies," *Computers and Education*, vol. 71, pp. 56–76, feb 2014.

[5] N.-T. Le, "A Classification of Adaptive Feedback in Educational Systems for Programming," *Systems*, vol. 4, no. 2, p. 22, 2016.

[6] S. Farid, R. Ahmad, and M. Alam, "A hierarchical model for e-learning implementation challenges using ahp." *Malaysian Journal of Computer Science*, vol. 28, no. 3, 2015.

[7] A. T. Bimba, N. Idris, A. Al-Hunaiyyan, R. B. Mahmud, and N. L. B. M. Shuib, "Adaptive feedback in computer-based learning environments: a review," *Adaptive Behavior*, vol. 25, no. 5, pp. 217–234, 2017.

[8] S. D'mello and A. Graesser, "Autotutor and affective autotutor: Learning by talking with cognitively and emotionally intelligent computers that talk back," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 2, no. 4, p. 23, 2012.

[9] A. Olney, S. D'Mello, N. Person, W. Cade, P. Hays, C. Williams, B. Lehman, and A. Graesser, "Guru: A computer tutor that models expert human tutors," in *Intelligent Tutoring Systems*. Springer, 2012, pp. 256–261.

[10] A. Gerdes, J. Jeuring, and B. Heeren, "An interactive functional programming tutor," in *Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education*. ACM, 2012, pp. 250–255.

[11] I. Arroyo, B. P. Woolf, W. Burelson, K. Muldner, D. Rai, and M. Tai, "A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect," *International Journal of Artificial Intelligence in Education*, vol. 24, no. 4, pp. 387–426, 2014.

[12] V. Rus, N. B. Niraula, and R. Banjade, "Deeptutor: An effective, online intelligent tutoring system that promotes deep learning." in *AAAI*, 2015, pp. 4294–4295.

[13] V. Rus, M. Conley, and A. Graesser, "The dendrogram model of instruction: On instructional strategies and their implementation in deeptutor," *Design Recommendations for Intelligent Tutoring Systems*, p. 311, 2014.

[14] V. Smy, M. Cahillane, and P. MacLean, "Sensemaking and metacognitive prompting in ill-structured problems," *The International Journal of Information and Learning Technology*, vol. 33, no. 3, pp. 186–199, 2016.

[15] M. Gawlowska, E. Beldzik, A. Domagalik, A. Gagol, T. Marek, and J. Mojsa-Kaja, "I don't want to miss a thing in learning dynamics and effects of feedback type and monetary incentive in a paired associate deterministic learning task," *Frontiers in psychology*, vol. 8, p. 935, 2017.

[16] S. Basu, G. Biswas, and J. S. Kinnebrew, "Learner modeling for adaptive scaffolding in a computational thinking-based science learning environment," *User Modeling and User-Adapted Interaction*, vol. 27, no. 1, pp. 5–53, 2017.

[17] C.-Y. Chou, K. R. Lai, P.-Y. Chao, S.-F. Tseng, and T.-Y. Liao, "A negotiation-based adaptive learning system for regulating help-seeking behaviors," *Computers & Education*, vol. 126, pp. 115–128, 2018.

[18] S. Hoch, F. Reinhold, B. Werner, J. Richter-Gebert, and K. Reiss, "Design and research potential of interactive textbooks: the case of fractions," *ZDM*, vol. 50, no. 5, pp. 839–848, 2018.

[19] M. E. Brock, M. A. Barczak, and S. A. Dueker, "Effects of delayed video-based feedback and observing feedback on paraprofessional implementation of evidence-based practices for students with severe disabilities," *Focus on Autism and Other Developmental Disabilities*, vol. 35, no. 3, pp. 153–164, 2020.

[20] T. Tuti, N. Winters, H. Edgcombe, N. Muinga, C. Wanyama, M. English, and C. Paton, "Evaluation of adaptive feedback in a smartphone-based game on health care providers' learning gain: Randomized controlled trial," *Journal of medical Internet research*, vol. 22, no. 7, p. e17100, 2020.

[21] J. Han, K. H. Kim, W. Rhee, and Y. H. Cho, "Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation," *Computers & Education*, vol. 163, p. 104041, 2020.

[22] A. T. Bimba, N. Idris, A. A. Al-Hunaiyyan, R. B. Mahmud, and N. L. B. M. Shuib, "Design of an algebraic concept operator for adaptive feedback in physics," in *Conference on e-Business, e-Services and e-Society*. Springer, 2018, pp. 181–190.

[23] A. Al-Hunaiyyan, A. T. Bimba, N. Idris, and S. Al-Sharhan, "A cognitive knowledge-based framework for social and metacognitive support in mobile learning." *Interdisciplinary Journal of Information, Knowledge & Management*, vol. 12, 2017.

[24] L. Gatti, M. Ulrich, and P. Seele, "Education for sustainable development through business simulation games: An exploratory study of sustainability gamification and its effects on students' learning outcomes," *Journal of Cleaner Production*, vol. 207, pp. 667–678, 2019.

[25] J. A. Ruiperez-Valiente, P. J. Munoz-Merino, and C. Delgado Kloos, "Improving the prediction of learning outcomes in educational platforms including higher level interaction indicators," *Expert Systems*, p. e12298, 2018.

[26] N. P. Moreno, A. D. Newell, L. R. Zientek, K. Nimon, and G. L. Vogt, "Linking science education and hiv using viral biology, epidemiology and science practices," *Health Education Journal*, p. 0017896918783778, 2018.

[27] S. Sanhueza Haro, A. Bravo Escobar, C. Faundez Araya, and E. Utreras Cofre, "Icts as cognitive tools of inclusion in physics class for high school students," *GONDOLA-ENSENANZA Y APRENDIZAJE DE LAS CIENCIAS*, vol. 13, no. 2, pp. 306–324, 2018.

[28] L. Hsu, "Measuring the effectiveness of summer intensive physics courses for gifted students: A pilot study and agenda for research," *Gifted Child Quarterly*, vol. 47, no. 3, pp. 212–218, 2003.

[29] K. Hochberg, J. Kuhn, and A. Maeller, "Using smartphones as experimental tools - effects on interest, curiosity, and learning in physics education," *Journal of Science Education and Technology*, pp. 1–19, 2018.

[30] C. Saini and J. Abraham, "Implementing facebook-based instructional approach in pre-service teacher education: An empirical investigation," *Computers & Education*, vol. 128, pp. 243–255, 2019.

[31] T. R. Johnson, "Violation of the homogeneity of regression slopes assumption in ancova for two-group pre-post designs: Tutorial on a modified johnson-neyman procedure," *The Quantitative Methods for Psychology*, vol. 12, no. 3, pp. 253–263, 2016.