

Affordances of Audience Response Systems: Effects of Instant and Regular Feedback

Ricardo Tamayo Osorio

Universidad Nacional de Colombia

Johanna Sánchez-Mora (✉ cijсанchezmo@unal.edu.co)

Universidad Nacional de Colombia <https://orcid.org/0000-0001-7333-4974>

Javier Corredor-Aristizábal

Universidad Nacional de Colombia

Research Article

Keywords: Audience Response Systems, University Lectures, Learning Monitoring, Feedback Density, Ecological Validity

Posted Date: May 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-442968/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Audience Response Systems (ARS), also known as clickers, are wireless devices commonly used in instruction. The present study explored the effects of ARS on students' performance in an introductory psychology course. The study also described the trajectories of students during the course. Participants in the experimental group used ARS to solve ten true/false questions throughout a three-hour lecture, held once a week. They received feedback immediately after providing an answer. The control group was exposed to the same questions in a paper-and-pencil format at the beginning of each lecture. They received feedback after seven days. The key dependent variable was quantitative performance on course quizzes and exams. Results show that students in the experimental group had statistically significant superior performance compared to the control group. Additionally, analysis of learning trajectories of students in both groups showed that the ARS group gradually progressed to higher performance, whereas the paper-and-pencil group maintained similar performance through the study. These results are discussed within the context of previous findings related to the effects of ARS on instruction. Particularly, we revisit research related to environmental affordances, learning monitoring, motivational factors, feedback density, and ecological validity.

Introduction

Audience Response Systems (ARS) are a powerful tool to enrich the experience of college education, given their potential to provide feedback to students. The effects of ARS have been extensively studied finding consistent effects in terms of motivation effects and mixed effects in terms of learning (Hunsu, Adesope & Bayly, 2016; Landrum, 2013; Rana & Dwivedi, 2016; Stowell & Nelson, 2007). These effects have been explained by ARS capacity to facilitate constant evaluation and student engagement. Studies in this area, however, have focused mainly on output measures, ignoring students' trajectories in learning experiences supported by ARS. Describing students' trajectories is core to understand the process that underlies learning in these contexts. The original contribution of this study is that, in addition to evaluating ARS effects in terms of output learning, from an educational point of view, it is important to have insights regarding the way in which students evolve when using ARS. ARS, beyond their apparent simplicity, entail a rupture with ingrained pedagogical logics, such as those that underlie the passive college lecture and the separation between learning and evaluation. In this logic, for instance, students listen passively during lectures and intervene only at discrete isolated points during the class. In the logic of ARS, on the contrary, evaluation happens constantly, and it is distributed throughout the lecture. ARS also imply logistical challenges such as organizing the delivery and collection of devices, which can be time-demanding and alter the normal routines of a college lecture. Both pedagogical gains and challenges are related to the affordances of ARS (Conole & Dyke, 2004). Affordances are defined as environmental features that permit and provoke behaviors and attunements during the learning process, which, in contemporary contexts, have been used to explain the pedagogical change fostered by technology. Analyzing students' trajectories might help to understand how they attune to the affordances

of ARS. Is it a linear trajectory of improvement? Is it immediate or gradual? Does it take time to get superior learning results using ARS?

An additional contribution of this article is that it evaluates the effects of ARS in a new socio-cultural context. It is important to study how different cultural and educational configurations affect learning with ARS. Research in this area has a limited geographical scope, focusing normally on first-world countries (Hunsu, Adesope, & Bayly, 2016). A broader geographical focus is necessary because ARS effects might depend on previously acquired student competencies, and on the study conditions and support-systems available in educational systems. Strong educational systems, such as those of developed countries, foster higher academic and metacognitive abilities, which determine to a large extent how students react to technology-supported teaching (Makransky et al., 2020). Similarly, educational conditions among countries vary and the effects of ARS might vary as well. Weaker support systems, such as large lectures without discussion sections, as it is usual in developing countries, might affect and diminish the effects of ARS. In this sense, there is a need to confirm and extend results previously found in first-world countries to the developing world. In consequence, this study explores the learning outcomes and trajectories of students attending a college lecture supported by ARS in a large city of a developing country. The theoretical framework and methodological approach are presented in the following pages.

Rationale for using ARS in college education

New technologies have the potential to enhance retention, accuracy, and transfer of information learned from university lectures (Blasco-Arcas, Buil, Hernández-Ortega, & Sese, 2013, Requena, 2008). Several strategies have been used to integrate new technologies in higher education. These strategies include fostering communities of learning using digital tools (e.g., Moodle) and incorporating authentic research activities supported by mobile devices into everyday classroom practices (e.g., Sun, 2014). Positive effects of these innovations have been documented using ethnographic methods, case studies, and quantitative measurements of various types (Graaff & Kolmos, 2003; Kolmos, 2004, Vargas, González, & Chávez, 2015). Authentic research activities and participative learning communities, however, have preconditions that are not necessarily present in every college lecture, such as small classroom size and sustained interaction between instructors and students. In contexts where these requirements are not met, technological innovations have been alternatively used to include problem solving and immediate feedback in instruction (Mayer et al., 2009). Passive listening in a lecture does not constitute an adequate pedagogical method because it does not allow students to apply their knowledge or review their learning process (Haidet, Morgan, O'malley, Moran, & Richards, 2004). Laboratory studies in highly controlled environments show that feedback contiguity and frequency are key factors for the identification of environmental regularities (Tamayo, 2003, 2015). However, traditional college education has not included teaching methods that allow students to monitor their own understanding during lectures (Bonvillian & Singer, 2013). In this context, students find few opportunities to transfer knowledge from theory to practice. One difficulty with alternatives to traditional lectures is that they require a high investment of instructors' time and attention.

In small classes, as opposed to large lectures, instructor-student interactions favor discussion and pedagogical support. In fact, previous studies have shown that students' perception of instruction quality is inversely proportional to classroom size (Gaviria & Hoyos, 2008). Additionally, small classrooms create conditions for frequent evaluation and constant feedback. Experimental and applied research has shown that frequent evaluation positively affects retention and application of theoretical knowledge. For instance, increasing test frequency before and after lectures improves students' long-term retention of complex topics irrespective of students' age or grade (Pashler et al., 2007; Roediger & Karpicke, 2006; Rohrer & Pashler, 2010). This fact suggests that evaluation itself can have positive intrinsic effects on memory by enhancing functional and contextualized recovery of contents and skills (Kromann, Jensen, & Ringsted, 2009; Mollborn & Hoekstra, 2010). It is clear then that presenting and solving problems within the classroom promotes effective learning and deep content understanding.

However, in practice, providing frequent feedback on test performance or problem-solving situations in large university lectures is time consuming and expensive (Mulryan-Kyne, 2010). One way to address this issue is the use of ARS, which consist of small devices ("clickers" or "keypads" in the US and "handsets" or "zappers" in the UK) wirelessly connected to a computer. Each student has an ARS device and sends an answer to a question or problem presented by the instructor. The computer integrates instantly all the answers and summarizes the results for the entire course, which can be then presented to the whole class, allowing feedback in real-time. In this context, the instructor can provide feedback to students and explain the logic behind the problems presented.

Effects of audience response systems

Previous research has shown that ARS improve interactivity in the classroom. Depending on the type of questions, they can increase students' attention and participation (Anthis, 2011). Although some researchers have found no improvement in final exams (Fallon & Forrest, 2011), they have reported an increase in students' responsibility to class assignments (Graham, Tripp, Seawright, & Joeckel, 2007; Jones, Crandall, Vogler, & Robinson, 2013). In particular, studies conducted with university populations have shown that ARS increase attendance and improve commitment to learning contents and subject matter, regardless of the difficulty level or academic field (Landrum, 2013; Rana & Dwivedi, 2016; Stowell & Nelson, 2007). Additionally, ARS can also favor critical thinking, support the application of concepts to real-life cases, and improve critical attitudes towards theoretical models (Mollborn & Hoekstra, 2010). Despite these potential benefits, students perceive that ARS generate anxiety when used to assess performance under a high-stakes grading scheme, especially in large groups (Nicol & Boyle, 2003). In other words, ARS have advantages, such as the opportunity to present problems and feedback in the classroom, but they also have disadvantages, such as increasing students' anxiety.

More generally, meta-analyses report that students' attention improves, especially when formative feedback is presented, and show positive effects in key motivational, attitudinal, and cognitive variables (Castillo-Manzano, Castro-Nuño, López-Valpuesta, Sans-Díaz, & Yñigues, 2016; Kay & LeSage, 2009). In particular, these results suggest that ARS have significant effects on knowledge transfer to new

situations, academic achievement, involvement, self-efficacy, and perception of lecture quality. However, fine-grained analyses show that the advantages of this type of technology vary depending on the comparison group. When compared to other strategies, which include frequent questions, ARS only have significant positive effects on motivational and attitudinal aspects, but not on academic achievement and learning. When compared to pedagogical strategies that do not include questions, ARS have a greater impact on learning (Hunsu, Adesope, & Bayly, 2016).

Explaining the effects of ARS

Positive learning effects of ARS can be explained by the joint influence of at least three specific factors: 1) active learning and involvement (2) constant questioning and evaluation, and (3) immediate feedback during pedagogical activities (Lantz & Stawiski, 2014). When using ARS, students move from passive listening to a pedagogical configuration in which they must actively respond to the instructor's questions, which, in turn, favors their involvement in the lecture. In fact, pedagogical research points out that these two factors – active learning and involvement – improve learning and academic performance (Chen, Gonyea, & Kuh, 2008). Similarly, questioning and evaluating students frequently has positive effects on retention and learning. Good questions afford higher cognitive processing and better concept integration (Rupp, Ferne, & Choi, 2006). Although ARS favor pedagogical dynamics that involve repeated questioning, this characteristic is not exclusive of this technology. Questioning strategies can be applied without ARS. Actually, a substantial body of research shows that ARS advantages over lectures are present only when they are compared to classes that do not involve questioning strategies (Hunsu, Adesope, & Bayly, 2016).

It is important to consider, however, that the characteristics of the pedagogical situation are not independent of tools used in the classroom. "Affordances" in ARS (Gibson, 1979/2014) facilitate and promote the use of pedagogical strategies such as questioning and providing feedback. The topic of affordances has been widely discussed in pedagogical research and is one of the cornerstones of conceptual analyses regarding the relationship between new technologies and pedagogical practices (Bower & Sturman, 2015). Although pedagogical strategies are clearly independent of the environment, this relation might be constrained by specific environmental properties involved in the learning context. For example, graphing calculators (or software for statistical analyses) may improve pedagogical practices because they carry out mechanical tasks, which helps students to focus on deeper pedagogical goals, such as abstraction and understanding. An important property of affordances is that learning requires attunement to environmental regularities of the environment and, specifically, to the tools and participation structures of the instructional situation (Conole & Dyke, 2004; Escallon et al., 2019). In educational research, this idea has been used to explain the trajectories of learning in educational experiences supported by technology (Dunleavy, Dede, & Mitchell, 2009).

The affordances of ARS may favor certain pedagogical practices. In particular, questioning teaching strategies might arise more frequently when using this type of devices. As tools designed to receive answers, ARS foster pedagogical strategies that include asking questions. Furthermore, ARS have affordances that amplify the effects of questioning, such as allowing immediate feedback on students'

responses. At a different level, ARS have advantages over alternative tools to present questions. Unlike strategies such as response cards or raising hands, they allow for students' anonymous answers, which facilitates participation, but provide the instructor with accurate information about students' opinions (Caldwell, 2007; Graham, Tripp, Seawright, & Joeckel, 2007). Additionally, frequent feedback seems to be a key factor to explain positive effects of ARS. Providing feedback helps error detection, which, in turn, helps students develop abilities to monitor their own learning process (Hattie & Timperley, 2007).

Whereas these factors suggest that ARS might have positive effects in college learning, research in this area is far from uniform. A substantial body of research has used qualitative descriptions of classroom experiences or end-of-course surveys (e.g., Castillo-Manzano et al., 2016; Fies & Marshall, 2006). In cases where systematic comparisons between ARS and other questioning strategies have been conducted, the contrast groups and methodology vary widely. For instance, in some cases, comparisons have been made using response cards (Desrochers & Shelnutt, 2012), mobile devices (Sun, 2014), color palettes (Brady, Seli, & Rosenthal, 2013) or simply asking students to raise their hands (Fernández-Alemán, Sánchez, López, & López, 2014). This diversity in comparison groups does not seem to address a key difference between ARS, on the one hand, and traditional assessment practices, on the other. If advantages linked to ARS arise from constantly evaluating and questioning students, then other pedagogical strategies involving these two variables would produce similar effects. If ARS afford particular teaching conditions beyond evaluating and questioning, then particular effects should be found when using ARS.

A suitable research strategy to solve this question is to use a comparison group that represents traditional learning conditions and includes also questioning and evaluation. A common practice in university education is partial evaluation through quizzes (Desouza & Flemming, 2003). Quizzes are used to motivate students to read class documents and to assess students' knowledge in between major exams (e.g., between mid-terms and finals). They also help maintain sustained attention and promote constant study. Therefore, if the effects of ARS rely solely on their ability to generate questioning conditions in the classroom, then there should be no differences between ARS and standard paper-and-pencil (P&P) quizzes. If ARS amplify the effects of questioning practices by allowing instant and distributed questioning and feedback throughout the lecture, then there should be differences between the two instructional methods. Knowing answers immediately after questions allows students to review their problem-solving strategy with a clear record of the cognitive process that led them to choose an answer. The cost, in this case, is that students in the ARS condition have to verify on their own whether or not their answers were correct. In the traditional P&P quiz version, students receive specific but delayed feedback, since the answer is recorded on paper and graded by the instructor after the session.

Another difference between the two options (P&P vs ARS) is distributed feedback throughout the lecture. Traditionally, quizzes are solved at the beginning of the lecture, whereas questions presented with ARS can be distributed throughout the whole lecture. This difference makes it easier for students using ARS to connect answers and contents, when memory traces are still recent. ARS provide real-time information to the instructor regarding students' understanding of his or her explanations, which allows the instructor to

adaptively correct comprehension errors. These facts have important theoretical implications. If ARS' advantages come only from evaluation and feedback, then quizzes should be as effective as ARS in improving performance. If, on the other hand, the advantages of ARS come from their affordances, which help students to monitor their reasoning process, provide instructors with real-time feedback about students' understanding, and distribute questioning throughout the lecture, then ARS must produce superior learning results than quizzes. Given that the primary interest of the present study is to compare these two pedagogical strategies (ARS and P&P) in an ecologically valid context, no attempt was made to separate the effects of immediate feedback, of distributed feedback throughout the lecture, and of information provided to the instructor about the students' responses.

Present Study

In the present study, we compare performance in the course exams under two different pedagogical strategies: (1) instant and distributed feedback provided by ARS; (2) traditional use of feedback provided by grading P&P quizzes (usually one week after completion). This design allowed us to evaluate whether a college lecture supported by ARS produces better learning results than a lecture using a traditional teaching strategy based on P&P quizzes. We also observe students' trajectories in both groups in the course quizzes throughout the term. The article's main contribution is that it evaluates the effects of ARS in a developing country with social and learning conditions that differ from those of prior studies. The article also describes the learning trajectories of students when attuning to the affordances of ARS, which opens a new line of research in the area with important implications for the design of learning experiences supported by clickers. Specifically, changes in quizzes results are described. In order to make performance in quizzes comparable the same 8 quizzes were applied to both groups. A description of students' trajectories is useful for understanding how participation evolves in two activity systems: one that is embedded in traditional P&P evaluation practices and another that is supported by the affordances of ARS. The specific details of this methodological strategy are presented in the following section.

Methods

This study uses a quantitative data analysis strategy to evaluate the effects of ARS on college learning, and to describe the trajectories of students in this process. To evaluate the effects of ARS, a t-test for independent samples was used with course exams' performance as the independent variable and condition (ARS vs P&P) as the independent one. To evaluate the trajectories a 2x8 mixed ANOVA designed was used, with condition as the between-subjects factor and quiz measure as the within-subjects factor. In this case, the performance in the quizzes was the dependent variable. In this study, a convenience, non-probability, sample was obtained in introductory psychology courses. The details of this method are presented below.

Participants

This study follows students during a 16-week semester in which they answered quizzes during 8 of these weeks and completed 2 exams. The core difference between the groups was that, in the P&P group, quizzes were presented at the beginning of the session, whereas in the ARS group quizzes were spread throughout the session. In this group, immediate feedback was provided, and answers were explained in the context of the topics contingently being presented by the instructor. This core difference was made possible by the affordances of ARS. The method used to collect these data is presented below. A total of 172 undergraduate students (age $M = 20.05$, $SD = 3.45$) initially enrolled in a course devoted to the philosophical foundations of psychology participated in the study. The course was introductory, and its contents were very specific to the epistemological debates underlying the historical development of psychology. No prior courses presenting the same contents were provided in the psychology coursework. Students were enrolled in two sections of the same course, occurring in different semesters. First, 123 students (44 female) were exposed to a traditional method of P&P quizzes. The following semester, in the ARS group, 47 students (17 female) were exposed to a teaching method of immediate feedback through ARS. The first group participated in a class in which traditional P&P quizzes were applied at the beginning of the lectures. The second group comprised students who participated in the same class taught in a different semester and in which the same quizzes were presented throughout the lecture using ARS. Both groups followed a lecture-only teaching strategy, in which the only contact with the instructor was the interaction during class. In particular, the only activities in which students engaged were listening to the lecture and answering quizzes and exams. This fact rules out the possibility of small group size increasing tutoring or permitting more detailed feedback. Similarly, students were instructed to work individually and were evaluated independently. This was done so to avoid any increase in interaction due to the reduced class size of the ARS group. Additionally, statistical controls were used in the data analysis to account for the group size difference. Two students abandoned the lecture and were excluded from further analyses. Given that the study was conducted in a natural educational setting, no randomization was possible and, therefore, the results of this study describe observational, not causal, patterns.

Materials

In the traditional instruction method that used P&P quizzes, students answered ten true/false questions about the assigned readings each week at the beginning of the lecture. Readings were assigned to students in the course's syllabus. Feedback about the accuracy of each response was provided to students one week later. In this case, students received information about whether or not their answers were correct. In the ARS method, the students answered the same ten true/false questions presented to the P&P quizzes group, but the questions were distributed throughout the lecture. Given that students in both groups answered exactly the same questions in each lecture, the core differences between groups were the distribution of questions throughout the lecture, the moment in which they were presented, and the fact that feedback was provided just after students answered the questions in the ARS group. These differences derived from the affordances of ARS systems and serve to illustrate the changes that clickers create in the pedagogical environment of a college lecture. In both groups, students received a grade according to how many correct answers they provided either through P&P quizzes or through the ARS system. The questions included both recalling factual knowledge and applying information in an

inferential fashion. They were designed by the first author and reviewed by other members of the research team to assure they connected with core theoretical elements of the course both at the memory and inference levels.

Students in the ARS group used S52Plus Audience Polling Devices developed by SunVote, working at a 2.4 GHz radiofrequency. Their size and weight enabled easy handling (92mm x 54mm x 8mm). Their weight was 32gm. ARS had a range of 30 meters which was adequate for the lecture halls in which classes were held. The keypad had both numbers, letters and erase and OK keys. When questions were presented during class, students were asked to click on the letter matching their answer. The "A" key and the "B" key corresponded to the true and false options and order was counterbalanced. ARS Questions were displayed in the lecture slides and they were introduced at moments during the lecture that connected with specific topics in the readings. In the P&P group, questions were presented at the beginning of the class in letter-size copies (Figure 1).

Procedure

We used a between-subjects design, in which each group was independently exposed to a different instructional method. However, the course contents (e.g., slides and lectures), the instructor, and the exams were the same. The course consisted of 16 face-to-face lectures of three hours, once a week. In eight of the 16 sessions, core contents of the course were presented. In those sessions, either the ARS or P&P quizzes were used, depending on the group. In the remaining sessions, the same activities were carried out for both groups. These activities included lectures without quiz, introductory sessions to the course and exams. There were not sample items or prior systematic training on the type of questions used in the ARS and P&P methodologies. Students had to figure out the use of the system as they were progressing in the course. Each question was introduced before each subtopic and students were informed immediately whether their answers were correct. For the two groups, the dependent variable was the students' average performance in two multiple-choice exams composed of 25 questions with four alternatives each (only one correct). For the two groups, the same contents and the same exams were used. The results of these exams were converted to a continuous scale ranging from 0 to 5 with an approximation to 1 decimal, 0 being the lowest grade and 5 the highest grade. The P&P group took the course one semester prior to the ARS group. In both semesters, students only could see the graded exams during the lecture and were not allowed to take pictures or notes of them. This procedure was implemented to prevent students in one semester to feed the answers to students in the following one.

In the ARS group, clickers were handed out at the beginning of the session, asking students in the class list to come to the front of the hall. They received the device in exchange for their student ID. At the end of the session, students returned the devices and received the ID back. Questions were presented in written format at specific points in the lectures' slides (Figure 1). Students had one minute to answer. After that, the instructor presented the answers chosen by students, the correct answer, and the rationale to choose that answer and discard the other options (Figure 2). No adjustments were made from students' answers and comment. In fact, the only difference between the slides in the P&P and ARS groups was the

introduction of quizzes' questions. Questions in the P&P group were presented as quizzes of ten questions at the beginning of the session. Feedback was given one-week later.

Results

Descriptive statistics show that the ARS group had a better performance than the P&P group (Table 1). This difference was observed both in the quizzes and in the exams. In the first case, the largest difference between both groups was found in the quizzes before the second exam (Quiz 2), which indicates a difference in the trajectories of both groups, probably associated with a gradual attunement to the characteristics and affordances of ARS. The exams show a similar pattern. Differences are small for Exam 1 and large for Exam 2. The standard deviations are similar in all groups, with slightly larger values in Exam 2 for the ARS group.

Mean differences between groups

In this analysis, the dependent variable was the average scores in Exam 1 and Exam 2 for each group, excluding any bonuses established in the syllabus (which were the same for the ARS or P&P groups). Levene's test indicated that variances were homogeneous in the two groups ($p = .996$). Therefore, a bilateral t -test for independent samples was conducted. This analysis showed significant differences between the two groups ($t(168) = 2.25, p = .02, d = .39, 95\% \text{ CI } (.03 - .51)$). Please note that t -tests and ANOVA are robust to group size differences when samples have homogeneous variances and sample sizes are relatively large (Sawilowsky & Blair, 1992). However, to control for possible statistical artefacts due to different sample sizes, the Brown-Forsythe and Welch corrections were calculated, showing significant differences between the P&P group and the ARS group on the this measure $F(1, 89) = 5.39, p = .02$. Both corrections were conducted to control by any source of bias due to the different sample sizes of the groups, although prior literature indicates that group sample size is only a problem when combined with different group variances (Skidmore & Thompson, 2013), which was not the case for this study. In any case, Welch's and Brown-Forsythe's corrections make the estimate of group differences more robust to violations of the variance homogeneity assumption (Lix, Keselman & Keselman, 1996; Reed III & Stark, 1988). ANOVAs and t -tests are built on three basic assumptions: independence of observations, normal distribution of error terms, and homogeneity of variances (Chen & Zhu, 2001), so equal sample sizes are not a requirement of these tests. Figure 3 shows the mean scores for each group.

Trajectories

Descriptive results suggest differences in the trajectories of both groups, with the ARS group being slightly superior before the first exam and highly superior after that. To further explore this question, a mixed General Linear Model (GLM) was built using the scores in quizzes as the within-subject factor and the groups as the between-subject factor. Results show a main effect of time indicating that the overall trajectories of both groups change in time $F(7, 1169) = 23.54, p < .00$. Results also show a significant effect of the interaction between time and condition, indicating that trajectories differ between groups $F(7, 1169) = 19.78, p < .00$. The Mauchly's test indicated that the sphericity assumption was not fulfilled

($W = .58, p < .00$), so the Greenhouse-Geisser and the Huynh-Feldt corrections were calculated. In all cases, the same results were obtained for both the overall trajectories and the interactions ($p < .00$).

As shown in Figure 4, the scores in the P&P group start higher than the scores in the ARS group, but end up at approximately the same level at which they started (around 3.5 points on a scale from 1 to 5). Scores in the ARS group start low, around 2.0 points, and then increase gradually to the 4.5 mark. Scores in the P&P group were similar at the beginning and at the end of the term, showing no significant improvement. When we tried to characterize these trajectories (Table 2), we found they were different for both groups. The model that better explained the trajectories of the ARS group was the cubic one ($p < .01$), although the linear and quadratic models were also significant. On the other hand, the P&P group trajectory could be only characterized as cubic with the same level of significance ($p < .01$) and as quadratic with a lower level ($p < .05$). The main difference is that the ARS trajectory can be characterized as linear whereas the P&P trajectory cannot. This is so because, unlike the ARS group, the P&P group did not show an increasing pattern during the study, which produces a zero slope in a linear model.

Discussion And Conclusions

This study showed that ARS has advantages over similar questioning strategies in the classroom. In particular, our study shows that students who used ARS scored higher in their course exams than students that solved the same questions in a P&P format. These effects can be interpreted as the combined result of three major factors. First, immediate feedback allowed students to review their answers, monitor their learning process, and inspect the logic underlying correct responses. Second, the results might be linked to the fact that questions were homogeneously distributed during the lecture, instead of grouped at the beginning of sessions. Organizationally, this effect would be difficult to achieve with P&P tests or other strategies because distributing questions and grading answers will interrupt the regular flow of class sessions. Distributing questions along the lecture favors higher attentional levels and allows students to connect quizzes' feedback to specific information. Third, positive effects might also relate to the fact that ARS enables instructors to gain information about students' ongoing understanding and can help them adapt their teaching process to the responses provided by students, thus correcting systematic misunderstandings that would otherwise go unnoticed.

Our primary goal was not to disentangle the effects of these three factors, but to explore their accumulated effect on the learning process. First, separate effects for each factor have been previously documented in many studies that do not necessarily reproduce the ecological conditions of college education (Janiszewski, Noel & Sawery, 2003; Kosba, Dimitrova & Boyle, 2007; Mayer et al., 2009). Our goal was, on the contrary, to evaluate the combined effects of two instructional strategies in an authentic teaching situation with high ecological validity. Although debatable, a large strand of educational research agrees that decomposing intervention factors is impossible in practice and not relevant when building theories of teaching, since what seems required is understanding how different features interact to produce aggregated effects in the classroom (Cobb, Confrey, DiSessa, Lehrer & Schauble, 2003). Our strategy is justified because the affordances of educational tools cannot be easily separated. ARS have

characteristics that favor pedagogical practices which simultaneously involve the three elements mentioned above. One could try to separate them, but that would be to some degree contrary to the intrinsic properties of the tool. For example, one could present all the questions using clickers at the beginning of the class, but that would do nothing more than duplicate the traditional P&P quiz. Alternatively, one could use clickers during the class but withhold information from the instructor (or keep the answers hidden from the students), but, in this case, the study would not be making full use of ARS' affordances, therefore, creating an unauthentic pedagogical situation.

The same applies to randomization. In our study, different groups were separated into different instructional methods, following the criterion of using the same instructor and the same content (Song, Kalet & Plass, 2016). Splitting students randomly would have changed the natural dynamics of the class. For example, let us suppose that P&P quizzes are presented at the beginning of each lecture for half of the students (randomly selected), while the other half waits outside the classroom. Then, the lecture starts normally for both groups, and, after some time, the other half of students (those who had not answered the quiz at the beginning) are asked to solve the same quiz questions using ARS, while the other half leaves the classroom or closes their eyes. This situation would have done nothing but alter the ecological validity of the study by introducing multiple extraneous factors (e.g., noise, cheating) in the design. Therefore, this study, although observational in nature, adequately reflects the effects of clickers in educational environments.

One limitation of this study is that groups had different sample sizes. Although this difficulty can be, and indeed was, controlled in statistical terms given the robustness of the tests used, different lecture sizes can have different effects on learning patterns. However, the class dynamics were similar in both cases. That is, the class in both cases 1) followed the same lecture style, 2) were given by the same instructor 3) covered the same content, and 4) used the same exams. Since the smallest group in this study had a size that did not allow more personalized interaction (e.g., use of workshops or seminar dynamics), it is safe to assume that the instructional dynamics were similar in both groups. However, the results must be taken with caution. Non-observable effects can arise from this difference. For instance, smaller class size might produce a less noisy environment or a more intense informal interaction among students outside of the classroom.

Regarding the trajectories of both groups in quizzes' scores, results show that each group has a different pattern. The ARS group shows an upward trajectory, whereas the P&P group has a steady performance. We must be careful when interpreting of these results, given that not direct observation of students' clickers use was obtained. However, these patterns suggest that students adapt progressively to the use of clickers. The requirements of ARS seem simple but, as explained, they entailed logistical and pedagogical changes. These changes might have affected the quiz performance at the beginning of the class, which would explain why, only after some time, higher results started to appear. In conceptual terms, these results might indicate a gradual attunement to the affordances of ARS. That is, at the beginning of the class, logistical challenges might have obstructed the flow of the learning activity and students might have had difficulty paying attention to questions and feedback. As students got more

experience with clickers, this coordination might grow smoother. Furthermore, students in the ARS group might have started with unclear ideas about the advantages and demands of clickers and about the transformations that they entail for the pedagogical process. For instance, they might have not expected to receive immediate feedback or have thought that the ARS quizzes were not an important activity in the class. Additionally, they might have started the class with unclear pedagogical expectations regarding their role in the activity. For example, they might have assumed that, as in traditional lectures, students' role was a passive one, centered on listening to the instructor. With time, students might have established clearer expectations regarding the effects of ARS by, for example, assuming a more active role during the questioning process. These results suggest that the affordances of ARS do not act immediately and that some time is required for students to get accustomed to the pedagogical configuration derived from them. Traditional P&P quizzes do not require this attunement because students are already accustomed to this type of pedagogical configuration.

Beyond methodological aspects, this study has broader pedagogical and theoretical implications. It reinforces the idea that incorporating technology in the learning process has positive effects when adequately used (Corredor, Gaydos & Squire, 2014; Corredor & Rojas, 2016; Castro-García, Olarte & Corredor, 2016; Diaz, Nussbaum, Ñopo, Maldonado-Carreño & Corredor, 2015; Kirkwood & Price, 2014). This study shows that facilitating students' ability to monitor their own learning process and review their answers does have positive effects. In this sense, ARS are a realistic option to support quality teaching in higher education when personalized small classroom activities are not possible. When organizing university lectures, it is assumed that college students are primarily autonomous and self-regulated, capable of organizing their own learning process and solving the challenges involved in learning from a lecture. This view contradicts research showing that university students have problems articulating their learning processes (Rodriguez & Clariana, 2017). In particular, educational research shows that control over the learning process is mainly beneficial for students with high metacognitive abilities and good previous knowledge (Corredor, 2006; Gonzalez, Maytorena, & Fuentes, 2018; Scheiter & Gerjets, 2007). In this sense, the use of ARS is an alternative to support students' learning in college.

Incorporating ARS in college education, however, requires attending to students' trajectories and to the needs that arise from the gradual attunement of students to the affordances of ARS. This study shows that the evolution of students' performance when using of ARS is not linear, which suggests a process of gradual attunement to the affordances of these devices (Dunleavy, Dede, & Mitchell, 2009). In this sense, this article adds to the existent literature showing that the effects of technology on education are related to the affordances of the particular devices being used (Bower & Sturman, 2015; Conole & Dyke, 2004). Furthermore, this study suggests that attunement to affordances is a key process when using digital technology. For this reason, teachers and instructors must carefully prepare the orientation and initial introduction to ARS. Routines for logistical purposes need to be clearly stated and practiced before ARS-based activities are introduced. Similarly, the pedagogical logic of ARS must be presented to students, so they understand why clickers are being used and what their advantages are. No immediate results must be expected, and instructors need to plan for initial gaps between their expectations and students' performance. ARS introduction in college lecture must be executed under the premise that tools do not

replace pedagogy. The impact of this ARS comes from the capacity to modify traditional pedagogical practices and engage students in constant active participation during the lecture, which implies changing students self-perceived roles as learners (Noronha & Batista, 2020). The instructor roles, conversely, must evolve from content providers to questioning and feedback designers. That is, the use of ARS in the college context requires not only the provision of devices and infrastructure but also a revision of the pedagogical contract

References

- Anthis, K. (2011). Is it the clicker, or is it the question? Untangling the effects of student response system use. *Teaching of Psychology, 38*(3), 189–193. doi:10.1177/0098628311411895
- Blasco-Arcas, L., Buil, I., Hernández-Ortega, B., & Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Computers & Education, 62*, 102–110. doi:10.1016/j.compedu.2012.10.019
- Brady, M., Seli, H., & Rosenthal, J. (2013). “Clickers” and metacognition: a quasi-experimental comparative study about metacognitive self-regulation and use of electronic feedback devices. *Computers and Education, 65*, 56-63. doi:10.1016/j.compedu.2013.02.001
- Bonvillian, W. B., & Singer, S. R. (2013). The online challenge to higher education. *Issues in Science and Technology, 29*(4), 23-30.
- Bower, M., & Sturman, D. (2015). What are the educational affordances of wearable technologies?. *Computers & Education, 88*, 343-353. doi:10.1016/j.compedu.2015.07.013
- Caldwell, J. E. (2007). Clickers in the large classroom: current research and best-practice tips. *Life Sciences Education, 6*(1), 9-20. doi:10.1187/cbe.06-12-0205
- Castillo-Manzano, J. I., Castro-Nuño, M., López-Valpuesta, L., Sanz-Díaz, M. T., & Yñiguez, R. (2016). Measuring the effect of ARS on academic performance: A global meta-analysis. *Computers & Education, 96*, 109-121. doi:10.1016/j.compedu.2016.02.007
- Castro-García, D., Olarte, F. D., & Corredor, J. (2016). Tecnología para la comunicación y la solución de problemas en el aula. Efectos en el aprendizaje significativo[Technology for communication and solving problems in the classroom. Effects on meaningful learning]. *Digital Education Review, 30*, 207-219.
- Chen, P. D., Gonyea, R., & Kuh, G. (2008). Learning at a distance: engaged or not. *Journal of Online Education, 4*(3), Retrieved from <https://nsuworks.nova.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1058&context=innovate>
- Chen, A., & Zhu, W. (2001). Revisiting the assumptions for inferential statistical analyses: A conceptual guide. *Quest, 53*(4), 418-439.

- Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational researcher*, 32(1), 9-13. doi:10.3102/0013189X032001009
- Corredor, J. (2006). General and domain-specific influence of prior knowledge on setting of goals and content use in museum websites. *Computers & Education*, 47(2), 207-221. doi:10.1016/j.compedu.2004.10.010
- Corredor, J., Gaydos, M., Squire, K. (2014). Seeing change in time: Video games to teach about temporal change in scientific phenomena. *Journal of Science Education and Technology*, 23 (3), 324-343. doi:10.1007/s10956-013-9466-4
- Corredor, J., & Rojas, L. (2016). Narrative and Conceptual Expertise in Massively Multiplayer Online Role Playing Games. *International Journal of Gaming and Computer-Mediated Simulations*, 8(1), 44-67. doi:10.4018/IJGCMS.2016010104
- Desrochers, M. N., & Shelnutt, J. M. (2012). Effect of answer format and review method on college students learning. *Computers and Education*, 59, 946-951. doi:10.1016/j.compedu.2012.04.002
- Desouza, E., & Fleming, M. (2003). A comparison of in-class and online quizzes on student exam performance. *Journal of Computing in Higher Education*, 14(2), 121-134. doi:10.1007/BF02940941
- Díaz, A., Nussbaum, M., Ñopo, H., Maldonado-Carreño, C., & Corredor, J. (2015). Orchestration: Providing Teachers with Scaffolding to Address Curriculum Standards and Students' Pace of Learning. *Educational Technology & Society*, 18 (3), 226–239.
- Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology*, 18(1), 7-22.
- Fallon, M., & Forrest, S. L. (2011). High-Tech Versus Low-Tech Instructional Strategies: A Comparison of Clickers and Handheld Response Cards. *Teaching of Psychology*, 38(3), 194–198. doi:10.1177/0098628311411896
- Fernández-Alemán, J. L., Sánchez, A. B., López, M. J., & López, J. J. (2014). Examining the benefits of learning based on an audience response system when confronting emergency situations. *CIN: Computers, Informatics, Nursing*, 32(5), 207-213. doi:10.1097/CIN.0000000000000053
- Fies, C., & Marshall, J. (2006). Classroom response systems: A review of the literature determinants of the results. *Journal of Science Education and Technology*, 15(1), 101–109. doi:10.1007/s10956-006-0360-l
- Gaviria, A., & Hoyos, A. (2008). Determinantes de los resultados de las evaluaciones de profesores: el caso de la Universidad de los Andes. [Determinants of the results of teacher evaluations: the case of the Andes University]. *Desarrollo y Sociedad*, (61), 87–117.

- Gibson, J. J. (1979/2014). *The ecological approach to visual perception: Classic edition*. NY: *Psychology Press*.
- Gonzalez, D., Maytorena, M. A., & Fuentes, M. A. (2018). Locus de control y morosidad como predictores del ejercicio físico-deportivo en estudiantes universitarios. [Locus of control and procrastination as predictors of physical-sports exercise in university students]. *Revista Colombiana de Psicología*, 27(2), 15-30. doi:10.15446/rcp.v27n2.61483
- Graaff, E. D., & Kolmos, A. (2003). Characteristics of problem-based learning. *International Journal of Engineering Education*, 657–662.
- Graham, C. R., Tripp, T. R., Seawright, L., & Joeckel, G. (2007). Empowering or compelling reluctant participators using audience response systems. *Active Learning in Higher Education*, 8(3), 233–258. doi:10.1177/1469787407081885
- Conole, G., & Dyke, M. (2004). What are the affordances of information and communication technologies?. *Alt-J, Research in Learning Technology*, 12, 113-124
- Escallón, E., González, B. I., Peña, P. C., & Rozo-Parrado, L. J. (2019). Implicaciones educativas de la teoría sociocultural: el desarrollo de conceptos científicos en estudiantes bogotanos. *Revista Colombiana de Psicología*, 28, 81-98. <https://doi.org/10.15446/rcp.v28n1.68020>
- Haidet, P., Morgan, R. O., O'malley, K., Moran, B. J., & Richards, B. F. (2004). A controlled trial of active versus passive learning strategies in a large group setting. *Advances in Health Sciences Education*, 9(1), 15-27. doi:10.1023/B:AHSE.0000012213.62043.45
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81e112. doi:10.3102/003465430298487
- Hunsu, N. J., Adesope, O., & Bayly, D. J. (2016). A meta-analysis of the effects of audience response systems (clicker-based technologies) on cognition and affect. *Computers & Education*, 94, 102-119. doi:10.1016/j.compedu.2015.11.01
- Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research*, 30(1), 138-149. doi:10.1086/374692
- Jones, S. J., Crandall, J., Vogler, J. S., & Robinson, D. H. (2013). Classroom response systems facilitate student accountability, readiness, and learning. *Journal of Educational Computing Research*, 49(2), 155–171. doi:10.2190/EC.49.2.b
- Kay, R. H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education*, 53(3), 819–827. doi:10.1016/j.compedu.2009.05.001

- Kirkwood, A., & Price, L. (2014). Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review. *Learning, Media and Technology*, 39(1), 6-36. doi:10.1080/17439884.2013.770404
- Kolmos, A. (2004). Estrategias para desarrollar currículos basados en la formulación de problemas y organizados en base a proyectos. [Strategies to develop curriculum based on the formulation of problems and organized based on projects]. *Educar*, (33), 077-96. doi:10.5565/rev/educar.262
- Kosba, E., Dimitrova, V., & Boyle, R. (2007). Adaptive feedback generation to support teachers in web-based distance education. *User Modeling and User-Adapted Interaction*, 17(4), 379-413. doi:10.1007/s11257-007-9031-z
- Kromann, C. B., Jensen, M. L., & Ringsted, C. (2009). The effect of testing on skills learning. *Medical Education*, 43(1), 21–27. doi: 10.1111/j.1365-2923.2008.03245.x
- Landrum, R. E. (2013). The Ubiquitous Clicker: SoTL Applications for Scientist-Educators. *Teaching of Psychology*, 40(2), 98–103. doi:10.1177/0098628312475028
- Lantz, M. E., & Stawiski, A. (2014). Effectiveness of clickers: Effect of feedback and the timing of questions on learning. *Computers in Human Behavior*, 31, 280-286.
- Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research*, 66(4), 579-619.
- Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2020). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*. Advance online publication. <https://doi.org/10.1037/edu0000473>
- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., Bulger, M., Campbell, J., Knight, A., & Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34(1), 51-57. doi:10.1016/j.cedpsych.2008.04.002
- Mollborn, S., & Hoekstra, A. (2010). "A meeting of minds": Using clickers for critical thinking and discussion in large sociology classes. *Teaching Sociology*, 38(1), 18–27. doi:10.1177/0092055X09353890
- Mulryan-Kyne, C. (2010). Teaching large classes at college and university level: Challenges and opportunities. *Teaching in Higher Education*, 15(2), 175-185. doi:10.1080/13562511003620001
- Nicol, D. J., & Boyle, J. T. (2003). Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education*, 28(4), 457–

473. doi:10.1080/0307507032000122297

Noronha, A. P. P., & Batista, H. H. V. (2020). Relações entre Forças de Caráter e Autorregulação Emocional em Universitários Brasileiros. *Revista Colombiana de Psicología*, 29, 73-86.

<https://doi.org/10.15446/rcp.v29n1.72960>

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing Instruction and Study to Improve Student Learning. IES Practice Guide. NCER 2007-2004*. National Center for Education Research. Retrieved from <http://eric.ed.gov/?id=ED498555>

Rana, N. P., & Dwivedi, Y. K. (2016). Using clickers in a large business class: Examining use behavior and satisfaction. *Journal of Marketing Education*, 38(1), 47-64. doi:10.1177/0273475315590660

Requena, S. R. H. (2008). El modelo constructivista con las nuevas tecnologías, aplicado en el proceso de aprendizaje. [The constructivist model and the new technologies, applied to the learning process]. *RUSC. Universities and Knowledge Society Journal*, 5(2), 6-. Retrieved from <http://dialnet.unirioja.es/servlet/articulo?codigo=2799725>

Reed III, J. F., & Stark, D. B. (1988). Robust alternatives to traditional analysis of variance: Welch W*, James JI*, James JII*, Brown-Forsythe BF*. *Computer Methods and Programs in Biomedicine*, 26(3), 233-237.

Rodriguez, A. & Clariana, M. (2017). Procrastinación en Estudiantes Universitarios: Su Relación con la Edad y el Curso Académico. [Procrastination in University Students: The Relationship to Age and Academic Year]. *Revista Colombiana de Psicología*, 26(1), 45-60. doi:10.15446/rcp.v26n1.53572

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. doi:10.1111/j.1467-9280.2006.01693.x

Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher*, 39(5), 406–412. doi:10.3102/0013189X10374770

Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading comprehension with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing*, 23(4), 441-474. doi:10.1191/0265532206lt337oa

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and Type II error properties of the t test to departures from population normality. *Psychological Bulletin*, 111(2), 352-360. doi:10.1037/0033-2909.111.2.352

Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review*, 19(3), 285-307. doi:10.1007/s10648-007-9046-3

Skidmore, S. T., & Thompson, B. (2013). Bias and precision of some classical ANOVA effect sizes when assumptions are violated. *Behavior Research Methods*, 45(2), 536-546.

Stowell, J. R., & Nelson, J. M. (2007). Benefits of electronic audience response systems on student participation, learning, and emotion. *Teaching of Psychology*, 34(4), 253–258.
doi:10.1080/00986280701700391

Song, H. S., Kalet, A. L., & Plass, J. L. (2016). Interplay of prior knowledge, self-regulation and motivation in complex multimedia learning environments. *Journal of Computer Assisted Learning*, 32(1), 31-50.
doi:10.1111/jcal.12117

Sun, J. C.-Y. (2014). Influence of polling technologies on student engagement: an analysis of student motivation, academic performance, and brainwave data. *Computers and Education*, 72, 80-89.
doi:10.1016/j.compedu.2013.10.010

Tamayo, R. (2015). Efectividad del reforzamiento inmediato en el aprendizaje de gramáticas artificiales. In Estudios sobre Comportamiento y Aplicaciones [Effectiveness of the immediate reinforcement in the learning of artificial grammars. In Studies on Behavior and Applications] (Felipe Cabrera, Óscar Zamora, Héctor Martínez, Pablo Covarrubias, Vladimir Orduña, Eds., Vol. IV, pp. 151–169). Ciudad de México, D.F.: UNAM. Retrieved from http://www.seminariosinca.org/docs/ESTUDIOS_vol4.pdf

Tamayo, R. (2003). Efectos del reforzamiento sobre el aprendizaje de secuencias generadas por una gramática artificial del estado finito. [Effects of reinforcement on learning of sequences generated by an artificial grammar of the finite state]. *Suma Psicológica*, 10(2), 211–222. doi:10.14349/sumapsi2003.176

Vargas, M. A. P., González, N. F., & Chávez, V. P. (2015). Efectos de un Entrenamiento Correctivo En La Elaboración Y Fundamentación De Preguntas De Investigación. Effects of Corrective Training in the Identification and Development of Research Questions]. *Acta Colombiana de Psicología*, 18(2), 87–100.
doi:10.14718/ACP.2015.18.2.8.

Tables

Table 1. Descriptive statistics for quizzes and exams. Mean (standard deviation in parenthesis).

Group	Quizzes 1	Quizzes 2	Total Quizzes	Exam 1	Exam 2	Total Exams
ARS	3.52	4.13	3.92	3.55	3.78	3.66
ARS	(.68) 3.48 (.51)	(.81) 4.25 (.53)	(.69) 3.87 (.40)	(.82) 3.54 (.67)	(1.03) 3.81 (.88)	(.84) 3.68 (.67)
P&P	3.47	3.74	3.60	3.38	3.42	3.40
	(.88) 3.47 (.89)	(.89) 3.74 (.98)	(.78) 3.61 (.78)	(.74) 3.39 (.74)	(.89) 3.42 (.89)	(.72) 3.40 (.72)

Table 2. Within-subjects contrasts for both groups.

	P&P Group		ARS Group	
Trajectory	F	Sig.	F	Sig.
Linear	1.38	.24	118.48	.00
Quadratic	5.37	.02	48.28	.00
Cubic	9.64	.00	34.99	.00

Figures

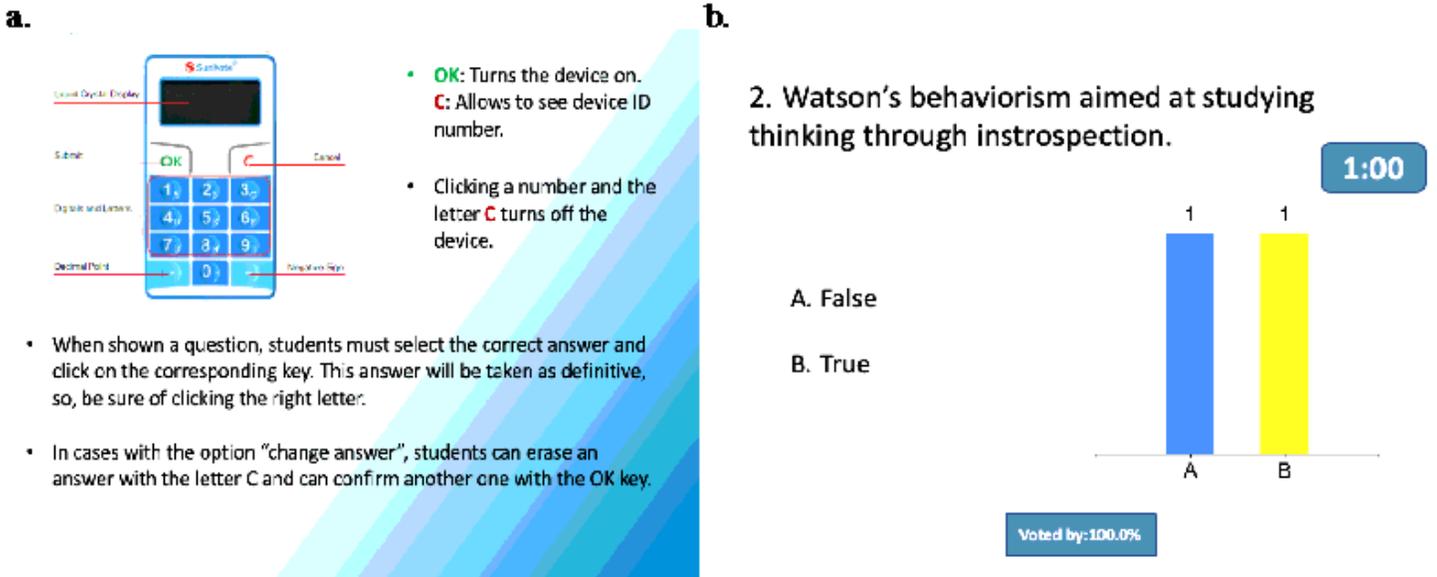


Figure 1

Examples of the slides used in the study. a) Introduction to the use of the device. b) Quiz question in the ARS group.

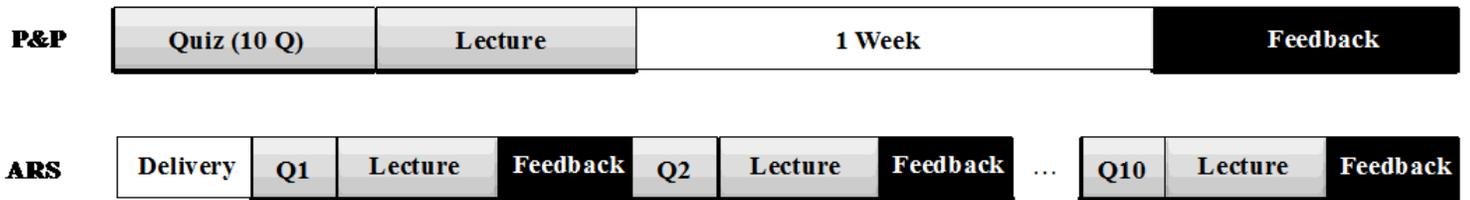


Figure 2

Procedure of study by group.

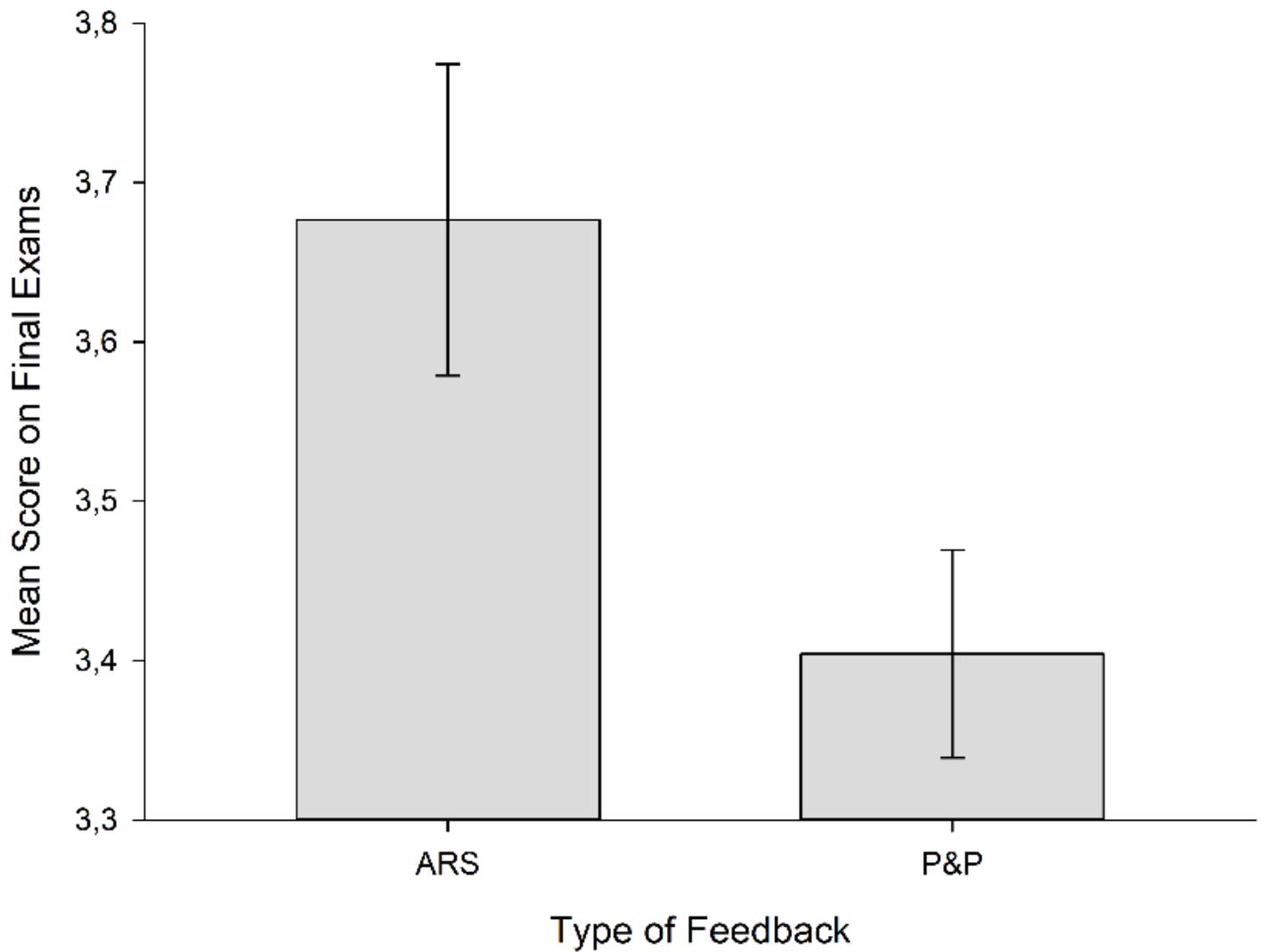


Figure 3

Mean scores for the Audience Response System (ARS) and the Paper and Pencil (P&P) feedback groups on the final exams. The score varies between 0 and 5. The error bars indicate +/- 1 Standard Error of the Mean (SEM).

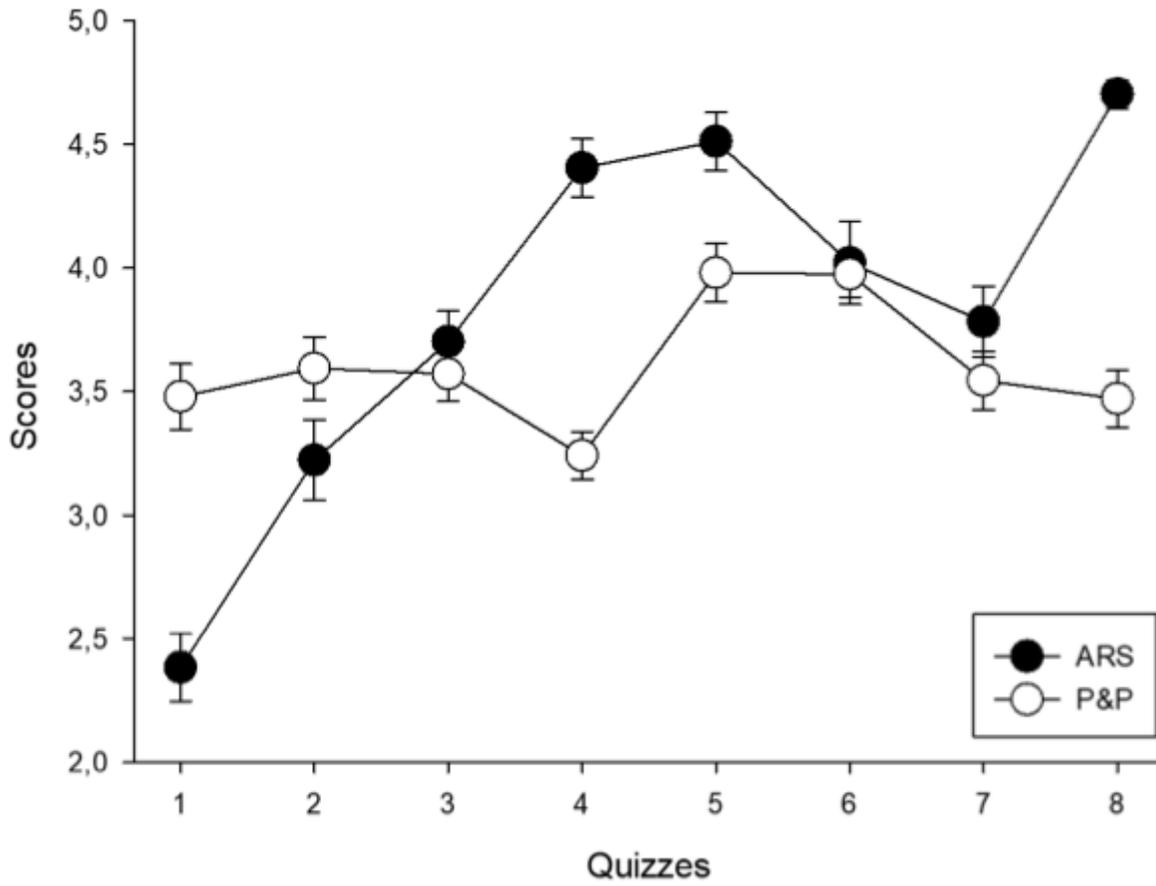


Figure 4

Quiz scores for the Audience Response System (ARS) and Paper and Pencil (P&P) feedback groups. The score varies between 0 and 5. The error bars indicate +/- 1 Standard Error of the Mean (SEM).