

ARTICLE

# A mixed-methods study of tailor-made animated cartoons in teaching punctuation in EFL writing

Arif Bakla

Cumhuriyet University, Turkey (arifbakla@gmail.com)

## Abstract

Despite their potential benefits, teacher-created animated cartoons have not found much room in second-language (L2) research, probably due to some technical challenges involved in creating them. This paper reports the findings of a mixed-methods embedded experimental study, designed to test the impact of tailor-made animated cartoons on the correct use of common punctuation rules in English. The participants were 112 Turkish-first language (L1) learners of English, assigned to either the treatment or control group through random cluster sampling. The instructional materials in the treatment group included teacher-created animated cartoons, exercises designed using SCORM-compliant software, and a forum for discussions in Moodle. The participants in the control group, on the other hand, used PowerPoint presentations (PPTs) instead of animated cartoons and completed the same follow-up activities. Quantitative results suggested that the treatment and control groups' post-test and late post-test scores significantly differed in favour of the former. Moreover, qualitative data from semi-structured interviews and document analysis revealed that the participants, especially those in the treatment group, viewed this learning experience highly positively. The findings globally imply that tailor-made animated cartoons might facilitate the learning of punctuation and help raise students' awareness of it.

**Keywords:** punctuation; animated cartoons; mixed methods; EFL writing; Moodle

## 1. Introduction

Punctuation has been one of the poorly investigated areas in second-language (L2) research (Hirvela, Nussbaum & Pierson, 2012; Waugh, 1998), and teachers often think that it is unimportant or spontaneously learnable. In fact, it is rather difficult to learn, particularly in L2 instruction (Alamin & Ahmed, 2012). As early as the 1950s, Coffin (1951) claimed that most college students had never learnt punctuation fully. Similarly, Backscheider (1972: 874) observed that it had been a “victim, scapegoat, and bogeyman for too many generations.” Hence, in contrast with other exciting issues in language teaching, punctuation seems rather humdrum (Gauthier, 1993). Today, the situation is perhaps worse as teachers often ignore it, and students leave punctuation marks out or pay little attention to them, due in part to the impact of mobile communication and the Internet. Cross-cultural differences with respect to punctuation rules also trigger misuse or total omission (Hirvela *et al.*, 2012). L2 learners may not be aware of the fact that L2 sentences are punctuated differently than those in their first language (L1). For example, in writing samples of Turkish-L1 learners of English, the impact of mother-tongue interference was apparent as students, particularly lower-level ones, usually used the rules in their

---

Cite this article: Bakla, A. (2019). A mixed-methods study of tailor-made animated cartoons in teaching punctuation in EFL writing. *ReCALL* 31(1): 75–91, doi:10.1017/S0958344018000046

L1 to punctuate L2 writing (Elkılıç, Han & Aydın, 2009; Kırkgöz, 2010). Similarly, another corpus study of undergraduate Turkish-L1 learners' writing concluded that the punctuation errors may be caused by L1 transfer or complexity of the connectives in English (Altunay, 2009). Moreover, lack of awareness of punctuation remains a major obstacle to accuracy in applying punctuation rules. Support for this comes from Benzer's (2010) study, in which the participants reported that they did not use punctuation, although they knew the rules. In conclusion, it seems rather challenging to teach or learn punctuation.

Researchers suggested the use of various methods to help learners accomplish this challenging task. For example, Coffin (1951) recommended the use of charts presenting a visual summary of punctuation rules. Another suggested method involved punctuating a text with no punctuation by listening to it and discussing deviances from the original text to raise students' awareness of punctuation (Gauthier, 1993). Other commonly used activities included error correction, proof-reading of peers' writing and discussing it, and so forth. Waugh (1998), for instance, recommended a discussion of punctuation in children's writing, group reading sessions, and instruction supported by exercises. Dawkins (2003) suggested teaching students how to use meaning for decision-making while punctuating sentences. Still other ways of teaching punctuation include, but are not limited to, the use of digital materials (charts, infographics, cartoons, and so on).

Although some preliminary work was carried out on teaching punctuation, there is need for more empirical research to test the effectiveness of innovative and potentially valuable tools. In this respect, with their tailor-made nature, teacher-created animated cartoons could help teachers who are up for designing effective instructional materials because, as Mayer and Moreno (2002) note, multimedia content has the potential to improve comprehension. It is now easier to produce animated cartoons, ACs for short, thanks to user-friendly software. In this respect, teachers can create their own ACs to explain complicated issues. However, this belief remains untested, as no one, to the best of the researcher's knowledge, has studied teacher-created ACs in L2 research. To test this, the researcher chose punctuation, as it has always been problematic according to not only previous research findings but also his hands-on experience.

As teacher-created ACs have not been studied so far, a discussion of the results of previous studies on commercially created ACs could inform this study and stimulate a deeper discussion of its findings. Therefore, the next section provides an overview of these studies following brief information on ACs.

## 2. Background to the study

### 2.1 *Animated cartoons*

Commercial cartoons, ACs, and authentic cartoons, as used in the literature, refer to animated movies produced for commercial purposes. They are movies used for entertainment (Mayer & Moreno, 2002). ACs have been used as authentic materials to teach various aspects of English (e.g. Arıkan & Taraf, 2010; Kristiansen, 2001; Meilleur, 2004). These movies are different from animations, as the term "animation" is generally understood to mean a set of pictures shown consecutively to describe or simulate the movement of objects.

### 2.2 *Studies on animated cartoons*

Given that animations are difficult to produce, previous research in applied linguistics exclusively focused on commercially created ACs. Most of these studies tested their impact on vocabulary acquisition and found that ACs facilitated L2 vocabulary acquisition (Alaba, 2014; Karakaş & Sarıçoban, 2012; Khodashenas, Farahani & Alishahi, 2014). In other studies, ACs were used to teach different language skills. For instance, Su and Liang (2014) reported that the use of ACs boosted the learners' motivation, yet they expressed a preference for watching cartoons in the same way they watched a movie rather than doing additional activities.

Likewise, Su and Liang (2015) found that the use of authentic ACs helped the participants develop their ability to match pictures with their descriptions. Such findings point to the effectiveness of ACs, but nearly 30% of the learners in Król's (2013) study reported that they were bored with studying vocabulary by watching them. This might justify Walker's (1999) warning that children may tend to watch videos passively (as cited in Król, 2013). These findings globally suggest that although watching ACs by following a lesson format could be quite useful, students may not like it.

ACs were also found to provide some affective benefits, including higher student motivation (Boucheix & Guignard, 2005; Malone, 1981; Su & Liang, 2014), increased attention (Moulic, 2012) and more positive attitudes towards English (Alaba, 2014). According to Moulic (2012: 74), learners "remain mentally, physically and emotionally motivated and ready to listen to the dialogues of the characters and stop attending to non-relevant activities as the film is in progress." This is because watching ACs involves an element of joy, so it is attractive, particularly for younger learners (Johnson, 2006; Khodashenas *et al.*, 2014). ACs break the routine and help decrease affective filter (Khodashenas *et al.*, 2014). In addition, they help boost the level of involvement (Lee, 2009). Teacher-created ACs might offer similar benefits, particularly if they are based on local needs and instructional design principles for multimedia are incorporated during creation. As design is a key issue in multimedia learning materials, the following section focuses on a commonly known theory as a framework that guided the researcher while creating the ACs used in the current study.

### 2.3 Cognitive theory of multimedia learning

Cognitive theory of multimedia learning, which provided support for the theoretical rationale of this study, is composed of two individual theories: dual coding theory (Paivio, 1991) and cognitive load theory (Chandler & Sweller, 1991). As a theory of memory and cognition, dual coding theory posits that information processing is carried out by the verbal system, which depends on words (written and spoken), and the nonverbal system, which utilises visual information (Mayer & Anderson, 1992; Paivio, 1991). According to Schnotz, Böckheler and Grzondziel (1999), these two subsystems interact with each other and visual information is processed by both verbal and visual systems. One of the key principles of this theory is that the learner forms various representational links between the incoming verbal information and the verbal representation of this new information and between pictorial information and the visual representation of this pictorial information. In addition, as the capacity of the short-term memory is limited, it might be easier to make referential links when verbal and visual information are adjacent (Mayer & Anderson, 1992; Mayer & Sims, 1994).

As for cognitive load theory, it suggests that information should be presented without overloading the cognitive system, due to memory constraints. Instructional materials are successful when learners are directed to primary rather than peripheral sources of data (Chandler & Sweller, 1991). Furthermore, students learn best when they are directed away from redundant information. If mutually referring information presented in two modalities is given in totally separate times and spaces, students have to integrate the information to make sense of it (Chandler & Sweller, 1991). This in turn increases cognitive load. In this sense, carefully designed and appropriately presented graphics help promote comprehension, learning, memory, communication, and inference (Tversky, Morrison & Betrancourt, 2002). Building upon these theoretical underpinnings, the next section presents an overview of the study.

### 2.4 Overview of the study

Today, animation tools make it possible for non-specialists to produce animations without much technical knowledge, and some projects have already made their way into instructional environments (e.g. DeCoursey, 2012; Stratton & Julien, 2014). However, teacher-created ACs are

relatively new, and there is little research on the use of user-friendly animation software in language teaching as an instructional design tool (DeCoursey, 2012). Furthermore, research on ACs has by and large addressed commercially produced cartoons with a focus on vocabulary or affective issues (Arikan & Taraf, 2010; Karakaş & Sariçoban, 2012; Khodashenas *et al.*, 2014; Król, 2013; Su & Liang, 2014; Su & Liang, 2015), whereas the present study focuses on punctuation as a problematic aspect of L2 writing. Unlike the previous research, which used commercially produced ACs, the present study sets out to investigate the effectiveness of teacher-prepared ones. Finally, most of the studies on this topic are quantitative in nature, but the researcher of the present study investigates the depth and breadth of the issue by using a mixed-methods research design. In brief, this study provides insight into how tailor-made ACs fit in English as a foreign language (EFL) writing instruction. To this end, it aims to address the following quantitative, qualitative, and mixed research questions:

- Q1. How do the scores of the participants in the whole study group change across the three testing times (pretest, posttest, and late posttest)?
- Q2. How does receiving instruction using tailor-made ACs affect the participants' learning and retention of problematic punctuation rules?
- Q3. How do the participants in the treatment group evaluate their experience of learning punctuation with ACs?
- Q4. How do the participants in the treatment group differ from those in the control group in their evaluation of the learning experience?
- Q5. How do the qualitative findings help us understand the nature of the experiment and its results?

### 3. Method

#### 3.1 Participants and setting

The participants in this study were 112 pre-intermediate Turkish-L1 learners of English (18 males and 94 females) registered for an EFL writing course aiming to teach English-major undergraduate students the basics of academic paragraph writing. The age of the participants ranged between 18 and 25. These students were required to study English intensively for a year before they could attend the classes in their department. They were assigned to the treatment or control group through random cluster sampling.

The data from the quantitative phase of the study informed the purposeful selection of the interview participants ( $n = 6$ ). The six qualitative interview participants were recruited based on three criteria: gender, posttest scores, and digital learning experience. The learners who reported that they had studied English using computers and the Internet were considered "more experienced," whereas those with limited or no experience of this were labelled as "less experienced." These labels were assigned based on self-report data collected using a mini survey administered after the treatment.

#### 3.2 Research design

This study adopted an embedded experimental mixed-methods research methodology. The researcher embedded qualitative data collection within a quantitative experiment (Creswell, 2012; Creswell & Plano Clark, 2011) to test the effect of tailor-made ACs on the learning and retention of basic punctuation rules and to gain insight into the perspectives of the participants. The data were collected sequentially at three points, with an unequal weight, emphasis being on the quantitative side: [QUAN(qual): QUAN = pretest, QUAN = posttest and survey, qual = semi-structured interviews and QUAN = late posttest] (see Figure 1). The researcher drew conclusions based on the data from both strands to understand the extent to which "the qualitative process

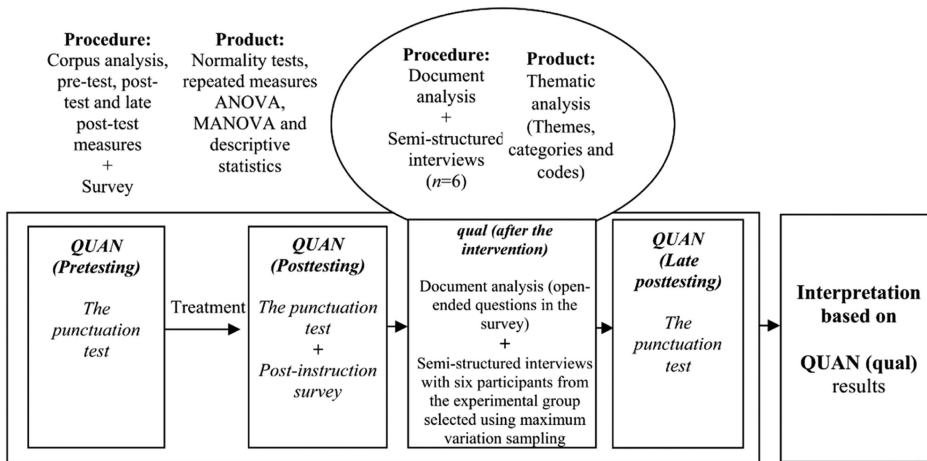


Figure 1. The procedural diagram illustrating the research design (adapted from Creswell & Plano Clark, 2011: 126)

Table 1. The results of the corpus analysis

Point of analysis	Usage	Wrongly punctuated (Frequency)	Wrongly punctuated (%)
Punctuation of conjunctive adverbs	174.00	68.00 <sup>a</sup>	39.08
Punctuation of coordinating conjunctions	171.00	84.00	49.12
Punctuation of essential/non-essential adjectival clauses	13.00	3.00	23.08
Punctuation of introductory phrases	70.00	21.00	30.00
Punctuation of quoted language	10.00	10.00	100.00
Punctuation of subordinating conjunctions	121.00	84.00 <sup>b</sup>	38.02
Comma splice	–	13.00	–
Total	559.00	245.00	43.83

<sup>a</sup>19 instances of semicolon use after conjunctive adverbs and 48 instances of no comma after them.

<sup>b</sup>38 errors resulting from *because* used in a fragment.

findings enhance the understanding of the experimental outcomes” (Creswell & Plano Clark, 2011: 223).

### 3.3 The procedure

Before the treatment, a corpus analysis of learner writing guided the identification of major punctuation problems. The corpus was composed of 80 randomly selected paragraphs (8,450 words) written as a regular classroom task in the paragraph writing class in which this study was carried out. As the corpus was relatively small, error/usage ratio rather than the number of errors was used to identify the punctuation problems to avoid wrong conclusions. Each error was identified and labelled using NVivo 10 to calculate this ratio.

As the corpus analysis indicates (see Table 1), the participants often had problems with the use of the comma and semicolon. They usually left out the comma after conjunctive adverbs or inserted a semicolon instead. They also tended to form a fragment by using *because*. They sometimes connected main clauses by using a comma or failed to use a comma before

coordinating conjunctions combining two complete sentences. They usually did not use a comma with non-essential adjectival clauses either. The researcher ignored less frequently occurring problems. Similarly, although participants sometimes failed to use a comma after introductory phrases, it was disregarded as it rarely harms comprehension.

Then the researcher developed a punctuation test and validated it (section 3.5.1). After the validation, both the treatment and control groups were given the pretest. The researcher administered the punctuation test immediately after the instruction and again one month later. In addition, the participants were given a survey to collect data about demographic variables for qualitative sampling and to learn their perspectives of the learning experience.

### 3.4 Materials

In the present study, all the instruction took place in Moodle. The participants studied the materials in a computer laboratory during regular class hours in a paragraph writing class, for which they met three times a week for 2 hours. The researcher's instructional role was limited to designing the materials (ACs, PPTs, and SCORM exercises), registering the students on Moodle, monitoring the class, and solving possible technical problems.

The treatment group watched nine ACs created by the researcher himself using GoAnimate. The ACs lasted approximately 4 minutes each, and they included characters teaching and learning punctuation in simulated environments, such as a school, clinic, or business meeting. They also used humour in the form of personification. Examples of such humour included Mr Comma's visiting his psychiatrist as he was stressed due to being misused and Mr Stick's looking for an expert to assist him to punctuate some sentences. In addition, the characters spoke in English (except those in the silent summary video) thanks to the built-in text-to-speech tool in GoAnimate (see online Supplement A for the ACs).

The control group studied nine PPT presentations. The rationale for this decision was that PPTs are so frequently used that they are now a dominant element of everyday teaching practices. This means that this study compared an innovative tool with a traditional one. Punctuation rules were explained in the PPTs and examples were provided; sometimes there were reminders and warnings shown in colour. Punctuation marks were usually highlighted to grab the attention of the students (see online Supplement B for a sample PPT).

In addition, both groups completed 15 SCORM exercises, wrote sample sentences, and discussed problematic points in a forum. The exercises were designed using authoringware. They included such activities as combining clauses or sentences by using appropriate connectors and punctuation, choosing the right punctuation, and error correction. The exercises also provided the answers or some feedback (see online Supplement C for sample SCORM exercises).

### 3.5 Data collection

#### 3.5.1 Quantitative data collection tools

The researcher developed a punctuation test with 30 dichotomous items based on the corpus analysis, and the same test was administered as the pretest, posttest, and late posttest (see online Supplement D for the punctuation test). The test-takers were asked to decide whether each item was correctly punctuated and to correct the errors to minimise the chance factor that potentially reduces reliability. Furthermore, sentence-level correction items were considered more appropriate as it might prove difficult to prepare a reliable test by using discourse-level ones, as a unit of meaning larger than a sentence might lead to different (and potentially correct) interpretations, which could harm reliability.

Content validity of the test was checked by seeking the opinion of an experienced EFL writing instructor and incorporating the results of the corpus analysis. The researcher was interested in how the results of the corpus analysis and the hands-on experience of this expert fitted with each other. It was found that the problem areas were highly similar. To examine the reliability of the



punctuation test, it was piloted with a similar cohort of students ( $n = 36$ ). The Kuder–Richardson 20, a good measure of internal reliability of tests with dichotomous response categories, such as “yes/no” or “right/wrong” (Creswell, 2012: 161; Perry, 2005: 134; Porte, 2002: 237), was calculated to be .80. During item analysis, the items with a discrimination index between .20 and .30 were revised (four items), and those below .20 were replaced (six items). The mean score for the test was 15.61, and the mean item facility was .51, which meant that the test was of medium difficulty.

The researcher also collected data from both groups after the experiment by using a survey with questions about age, gender, and the extent of previous experience with digital learning. The survey also included an open-ended question seeking the participants’ perspectives of the negative and positive sides of this learning experience. As written responses to open-ended questions in surveys can be considered as documents (Patton, 2002: 4), the responses to this open-ended question were analysed within the qualitative strand. Forty-three participants from the experimental group and 41 from the control group volunteered to take the survey, so two randomly selected participants from the former were excluded from the document analysis.

### 3.5.2 *Qualitative data collection tools*

After the treatment, semi-structured interviews were carried out with six participants from the treatment group. During the interviews, the interview guide approach was used to obtain the participants’ reactions to learning punctuation by using ACs. There were 15 questions along with some probes, which were created based on the data from the literature, the researcher’s observations, and Moodle logs. The interview was piloted with two students, and the wording of some items was changed. The interview was conducted in Turkish, and the direct quotes presented in the results section were translated into English by the researcher himself. Another significant source of data was the participants’ responses to an open-ended question included in the survey.

## 3.6 *Data analysis*

### 3.6.1 *Quantitative data analysis*

The first step to data analysis was the analysis of a small-scale corpus of learner writing. The errors in the corpus were labelled using NVivo 10 and some descriptive statistics were calculated. SPSS 18.0 was used to carry out tests of statistical significance on punctuation test scores. Before carrying out such tests, the researcher checked assumptions of each test. To carry out within-subjects comparisons, a repeated measures ANOVA was used. The alpha level ( $\alpha$ ) was set to .05 in all of these inferential statistical analyses, except for the MANOVA in which significance level was divided by 3 ( $.05/3 = .0167$ ). The qualitative data analysis was carried out using NVivo 10.

### 3.6.2 *Qualitative data analysis*

In the qualitative strand, the researcher prepared a flexible plan to analyse the data. He followed a three-phase approach to qualitative coding introduced by Strauss and Corbin (1998). The first step was to examine the transcripts before coding. A list of codes was created by using the data from the literature, quantitative results, and a preliminary analysis of the transcripts. The data were coded using these constructed codes and a few in vivo codes. To ensure the internal and external integrity of the data, the categories were checked to ensure homogeneity in a category and heterogeneity across categories (Given, 2008; Patton, 2002). At the initial stages of coding, the questions asked to aid the comprehension of the data guided the coding process and helped spot ideas. Analytic memos and annotations helped the researcher record new ideas, create relationships between codes/categories, and provided food for thought in subsequent cycles of coding. The researcher used negative case analysis and disconfirming evidence to enrich the discussion, reinforce the data, test the ideas, and revise them (Given, 2008; Patton, 2002; Tracy, 2013). Coding went on until the saturation point. Also, queries in NVivo were used to further

**Table 2.** Descriptive statistics about the groups' scores across the three testing times

	Group	<i>N</i>	<i>M</i>	<i>SD</i>	Range
Pretest	Experimental	62	13.85	4.35	21
	Control	48	13.75	3.72	17
	Total	110	13.81	4.07	21
Posttest	Experimental	62	20.69	5.18	21
	Control	48	18.17	4.82	20
	Total	110	19.59	5.16	22
Late posttest	Experimental	62	21.45	4.45	19
	Control	48	18.77	5.44	22
	Total	110	20.28	5.06	22

explore and compare the data based on respondent characteristics (see online Supplement E for additional notes on the qualitative data analysis).

## 4. Results

### 4.1 Quantitative results

*Q1. How do the scores of the participants in the whole study group change across the three testing times?*

The participants' test scores improved from the pretest to posttest and almost remained unchanged in the late posttest (see Table 2). To test whether this change was significant, a one-way within-subjects repeated measures ANOVA was conducted to compare the effect of instruction on the participants' test scores across the three testing points.

The data were initially checked for normality. Two outliers were removed, yet the late posttest data were not normally distributed. However, this was not considered a serious problem, as ANOVA is a robust test against lack of normality that is not caused by outliers. In addition, error variances were homogeneous, but Mauchly's *W* test indicated that the assumption of sphericity had not been met,  $\chi^2(2) = .930$ ,  $p = .02 < .05$ . Therefore, the results of the multivariate tests, which do not require the sphericity assumption, were used. Being a robust test against the violation of assumptions, Pillai's trace test was used and it was significant,  $.687$ ,  $F(2, 107) = 117.59$ ,  $p = .001$ ,  $\eta_p^2 = .69$ , power = 1.0. This indicated that the mean scores significantly differed across pretest, posttest, and late posttest. The effect size of the change across the time points was relatively large.

Pairwise comparisons using the Bonferroni correction revealed that the instruction elicited a significant increase in the participants' test scores from the pretest ( $M = 13.81$ ,  $SD = 4.07$ ) to posttest ( $M = 19.59$ ,  $SD = 5.16$ ), which was statistically significant ( $p = .001 < .05$ ). On the other hand, their mean scores remained almost the same in the late posttest ( $M = 20.28$ ,  $SD = 5.06$ ) with their posttest scores ( $M = 19.59$ ,  $SD = 5.16$ ) (see Figure 2), and the difference was not statistically significant ( $p = .15 > .05$ ). Therefore, it was concluded that instructional practice helped boost the participants' scores at statistically significant levels, and the improvement remained unchanged one month later.



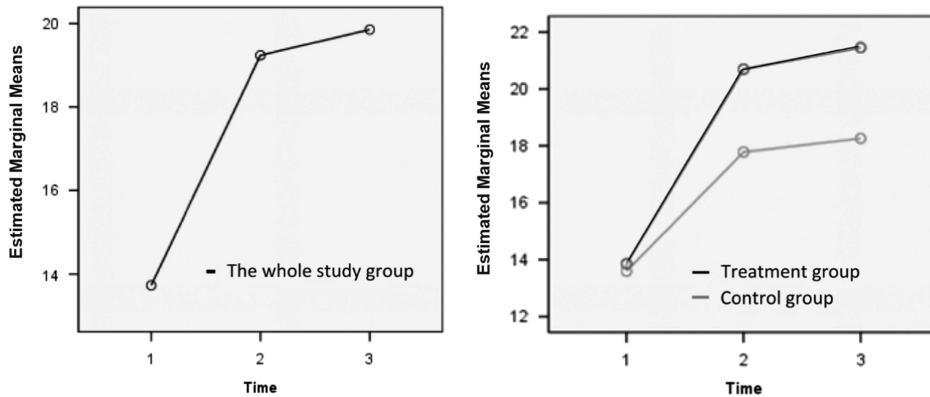


Figure 2. Within-subject comparisons across the three time points. Time 1=the pretest, Time 2=the posttest and Time 3=the late posttest

Table 3. Results of the multivariate analysis of variance

Test	$F$	$p$	$\eta_p^2$	Power
Pretest	.018	.894	.000	.052
Posttest	6.839	.010	.060	.736
Late posttest	8.077	.005	.070	.804

To investigate the longitudinal differences between the pretest, posttest, and late posttest between the experimental and control group, the interaction effect (time\*group) in the repeated measures ANOVA was examined. Pillai's trace test was used, and it was significant, .088,  $F(2, 107) = 5.13$ ,  $p = .007$ ,  $\eta_p^2 = .088$ , power = .81. This finding indicates that the longitudinal change between the groups is significantly different, although the effect size was moderate (Cohen, 1988).

Q2. *How does receiving instruction using tailor-made ACs affect the participants' learning and retention of problematic punctuation rules?*

The results of the repeated measures ANOVA indicated that there was a longitudinal change, yet these results did not indicate if the mean scores of the experimental and control group differed across the three testing points. A one-way MANOVA was carried out to investigate this. Other assumptions (besides the ones carried out for the repeated measures ANOVA) were also checked before carrying out the MANOVA. Maximum Mahalanobis distance ( $D^2 = 14.05$ ) was within the allowable limits.

As MANOVA is a robust test with no non-parametric equivalent, it was carried out despite minor violations of its assumptions. The test revealed a significant multivariate main effect for type of instruction (ACs versus PPTs) on the test scores, Pillai's trace = .101,  $F(3.00, 106.00) = 3.985$ ,  $p = .010$ ,  $\eta_p^2 = .101$ , a medium effect size. Power to detect the effect was .824. Given the significance of the overall test, the univariate main effects were examined. No significant univariate main effects for type of instruction were obtained for the pretest,  $F(1, 108) = .018$ ,  $p = .894$ . On the other hand, univariate main effects were significant for the posttest,  $F(1, 108) = 6.839$ ,  $p = .010$ ,  $\eta_p^2 = .060$ , power = .736, and the late posttest,  $F(1, 108) = 8.077$ ,  $p < .005$ ,  $\eta_p^2 = .070$ , power = .804. The effect size was moderate in both cases (see Table 3).

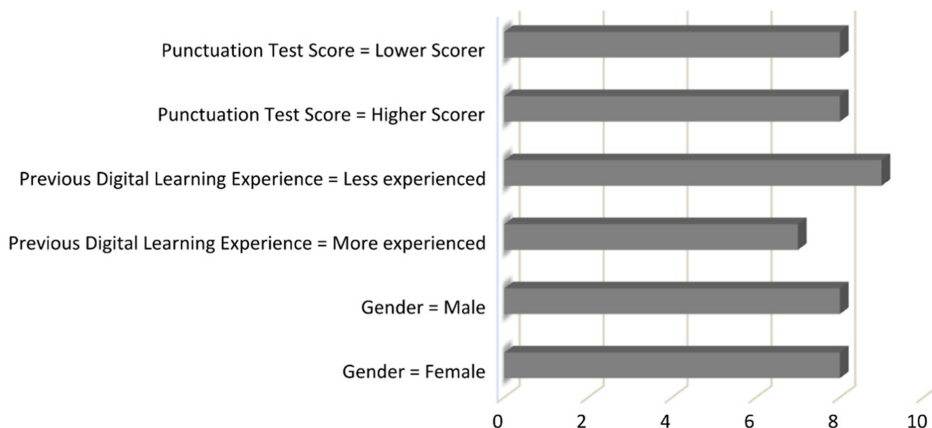


Figure 3. The distribution of the positive comments from the six respondents (each adjacent pair of bars represents the 16 positive comments)

No significant pairwise differences were obtained in the pretest in the mean for the experimental group ( $M = 13.85, SD = 4.35$ ) and control group ( $M = 13.75, SD = 3.72$ ). This indicated that their mean scores were roughly equal before the experiment. However, a significant mean difference between the experimental ( $M = 20.69, SD = 5.18$ ) and control group ( $M = 18.17, SD = 4.82$ ) was obtained in the posttest scores. There was a statistically significant difference between scores of the experimental group ( $M = 21.45, SD = 4.45$ ) and control group ( $M = 18.77, SD = 5.44$ ) in the late posttest as well. These results from the repeated measures ANOVA and MANOVA globally indicate that both groups improved their posttest scores at significant levels, and they retained their knowledge in the late posttest. However, the instruction provided in the treatment group was more effective.

### 4.2 Qualitative results

#### 4.2.1 Data from the qualitative interviews

The analysis of the qualitative interviews revealed three main themes. The contents of these themes globally answered the following research question:

Q3. How do the participants in the treatment group evaluate their experience of learning punctuation with ACs?

*Theme 1: The overall experience appeared mostly positive.*

As is apparent in Figure 3, the respondents adopted a positive attitude towards this learning experience (six participants making 18 positive references). Interestingly, less experienced learners talked about the experience more positively (count = 3, reference = 9). One participant expressed his desire for similar learning experiences by saying, “I would like to be involved in as many activities like this as possible” (P05). Similarly, another said, “If only we could use this for all classes” (P04). Such a positive experience seemed to have transformed their attitude towards online learning. The following quote, for example, signalled this apparent transformation: “I had never done such exercises on the Web before ... In this respect, I have a highly positive idea of it. I think that we can receive instruction on the Web from now on” (P05).

In addition, this learning experience in the treatment group had some strong sides that most of the respondents stressed. For instance, some respondents reported that the ACs used both the audio and visual channel, so they helped them remember better (count = 3, reference = 4). The participants also thought that the ACs and accompanying exercises were

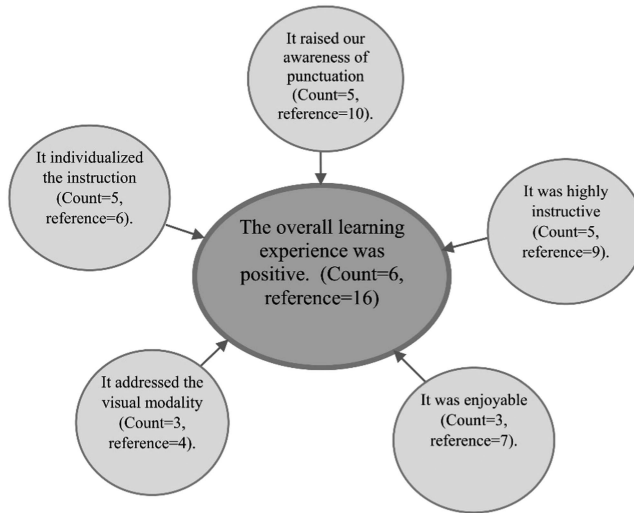


Figure 4. The strong sides of the learning experience

Table 4. Respondents’ view of punctuation before and after this digital learning experience

Participant	Before the learning experience	After the learning experience
P01	“As I said, I did not pay attention to punctuation or care about it. I used to ask myself, ‘How important could it be or how serious meaning problems could it cause?’”	“Then I began to pay attention to punctuation.”
P02	“Actually, I did not have any idea of how to use punctuation marks and conjunctions.”	“Well, my current knowledge of punctuation is far from perfect, yet think that I have decreased the number of mistakes I used to make.”
P03	...	“I recognised that I had made silly punctuation mistakes until now. From now on I know where to use punctuation marks.”
P04	“I did not know that punctuation marks were so important.”	“Now I have come to realise that punctuation is something important, so my perspective has changed ...”
P05	“I did not have information about punctuation marks ... I never paid attention to them because I was not interested in knowledge of punctuation.”	“After this experience, I gained lots of knowledge of punctuation.”

highly instructive. However, only one participant stressed the tailor-made nature of the ACs. She said, “You prepared the exercises yourself. You knew what we needed to study” (P04) (see Figure 4).

Another significant contribution of this learning experience was that the participants learnt a lot about punctuation and developed an awareness of it. Most of them reported that they had not been aware of the importance of punctuation or they had disregarded it. However, having realised its importance after this particular learning experience, they began to check it in their writing (see Table 4).

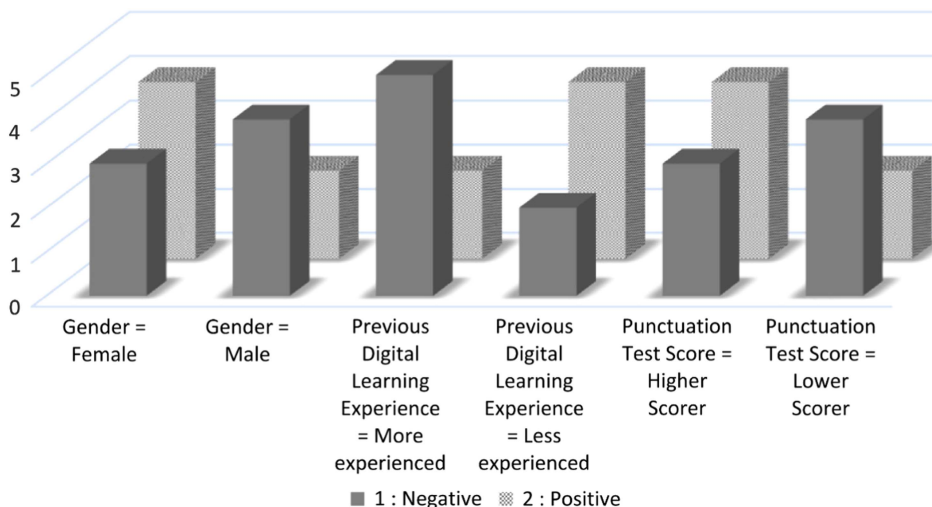


Figure 5. The participants’ perspectives on the individualization of the instruction

Despite predominantly positive ideas, some participants adopted different perspectives. For instance, P01 suggested that such online activities should be used to introduce variety into the classroom rather than substitute traditional learning. She said, “You see ... it is an alternative. It is wise to be open to innovations” (P01). Another student (P06) reported that she found ACs a little bit childish but not negative.

*Theme 2: The participants used some strategies while watching the ACs.*

Students mostly took down notes and some of them captured the computer screen, but one student reported that she did not take any notes or screen captures. Negative case analysis indicated that this participant considered it more like a film rather than an instructional video. She said, “I don’t like to pause the video to take down notes as I concentrate on the movie” (P01). This idea was shared by some others as well, but they watched the video nonstop in the first viewing and took down notes later.

*Theme 3: There were some problems involved in the process.*

One of the participants reported that she felt awkward as she did not see it as meaningful to watch the videos repetitively. Nonetheless, the participants were able to progress at their own rate thanks to the learning materials (count = 5, reference = 6). Henceforth referred to as individualization, this was mentioned with its both positive and negative sides. Initially, the researcher thought that individualization could appeal to most of the participants. However, the qualitative data indicated that the participants’ use of the term also connoted self-study, which was probably associated with a lack of self-confidence as they felt that they were unable to get involved in face-to-face communication with the teacher or peers. This feeling of a lack of self-confidence was apparent in the following quote: “When I am unable to communicate with the teacher directly, I feel like not having self-confidence. You must have the chance to ask the teacher your questions; it must be direct communication” (P01). Interestingly enough, digitally more experienced participants thought that individualization had some negative sides (see Figure 5).

Finally, although the Moodle forum was perceived as useful, some learners thought that it was not used efficiently by the students in general. Moreover, some students thought that writing sample sentences in the forum led to anxiety. For example, one participant said, “As others see what I write there, I think twice before writing. Therefore, I was late for participating in the

forum discussions” (P05). An examination of the forum posts confirmed this because only particular students participated in the forum discussions.

#### 4.2.2 Data from the document analysis

*Q4. How do the participants in the treatment group differ from those in the control group in their evaluation of the learning experience?*

The findings from the document analysis mostly confirmed those from the interviews and provided some comparative data by incorporating the perspectives of the control group. As with the interview participants, the respondents from the treatment group viewed their learning experience in a quintessentially positive way (count = 22, reference = 60). Similarly, those in the control group seemed to have enjoyed studying punctuation using PPTs and exercises, although the number of positive comments in this group was much smaller (count = 14, reference = 25). A marked difference between the comments of the two groups was that the number of comments on specific aspects of the learning experience in the treatment group was much higher. For example, 15 respondents reported that the learning materials, particularly the ACs, were highly instructive. They frequently used the adjectives “instructive” and “catchy” in their descriptions. However, only three respondents in the control group used such adjectives. Furthermore, more respondents in the treatment group than the control group found the materials enjoyable, interesting, motivating, and creative.

Despite predominantly positive comments, several students viewed the experience negatively. One respondent from the treatment group found the ACs distracting, while another found them not only boring, but also childish. Still another respondent thought that they were boring and the characters were unlikable. In addition to this, some students thought that they seemed more appropriate for children (count = 9, reference = 11); some of these people used the same adjective (childish) that one of the interview participants (P06) used to describe the ACs. Moreover, four students from the treatment group reported that they were bored either with repetitive watching or numerous exercises (count = 4, reference = 8). Although there were no negative comments about the global learning experience in the control group, the respondents found the PPTs boring and monotonous (count = 8, reference = 11).

The learners also suggested some improvements to the instruction. The suggestions in the experimental group predominantly concerned the voices of the characters (count = 35, reference = 65). Fifteen respondents stressed the robotic nature of the voices by using phrases that roughly meant “robotic but intelligible.” Some students also suggested that sound quality should be improved. However, as the sound was of HD quality, the word “quality” was probably used in an extended way to imply the use of human and age-appropriate voices for the characters. Finally, in line with the comment of P01, one respondent from each group suggested that online learning should be supported by additional face-to-face sessions.

These results globally suggest that the data from the qualitative interviews and document analysis seem to support each other. The data also indicate that the learning design was not bad for the control group, but a considerable number of learners suggested adding audio and/or video as learning materials (count = 17, reference = 17), and some were bored with the presentations. Fewer learners in the control group than the treatment group viewed the learning experience positively.

## 5. Discussion

This section presents a discussion of what the quantitative results mean and how the qualitative findings help us understand the nature of the experiment and its results. Both the experimental and control groups improved their scores in the posttest at significant levels, yet the improvement in the former was significantly higher. The improvement of the scores in both groups in the posttest might be attributed to the nature of the instruction with plentiful exercises, created based

on a corpus analysis of learner errors. The strong effect size for the change across the three tests in the whole study group ( $\eta_p^2 = .69$ ) not only supports this but also indicates that if appropriate materials are used, punctuation, a notoriously difficult subject for Turkish students, could be taught successfully. On the other hand, the superior performance of the treatment group in the posttest and late posttest can be attributed to the ACs, despite the moderate effect sizes for the posttest and late posttest. This is because, as Su and Liang (2015) concluded, evidence from qualitative data indicated that ACs were appealing for the learners.

Although statistical tests revealed moderately higher gains in the treatment group than in the control group in the posttest, the ACs were superior to the PPTs, mostly from a qualitative perspective. As noted earlier, half of the interview participants found it really enjoyable (7 references from three participants), and almost all of them thought that the materials were highly instructive (see Figure 4). According to the data from document analysis, they seemed more attractive than the PPTs as the learners in the experimental group talked about the learning experience much more positively than those in the control group (60 versus 25 positive comments respectively). It should also be noted that the positive comments from the respondents in the control group were related to the learning design in general and Moodle exercises rather than the PPTs. In short, more positive attitudes in the treatment group can be attributed to the use of the ACs instead of the PPTs, indicating that the ACs and exercises seemed to be a better combination than PPTs and exercises.

The results of the experiment and the qualitative data are consistent with the limited research available. They concur well with the results of Alaba's (2014) study, in which the use of ACs boosted the participants' score in the achievement test. The moderate success of the ACs in the present study also seem to support Mayer and Moreno's (2002) claim that the use of multimedia tools and verbal language boosts comprehension and retention. Evidence from the qualitative interviews supported this because half of the interview respondents noted that the ACs addressed the visual modality and helped improve their knowledge of the punctuation rules. Similarly, one survey respondent from the treatment group also mentioned this, whereas none in the control group made such a comment; rather, they voiced their desire to study with more audiovisual materials.

Another remarkable result to emerge was that tailor-made ACs contributed to the affective domain. That is, they helped the participants develop positive attitudes towards online learning and motivated them to study. Moodle logs also supported this because it was apparent that they watched the videos several times and completed all the exercises. This finding favourably correlates with those of various studies (Boucheix & Guignard, 2005; Khodashenas *et al.*, 2014; Malone, 1981; Su & Liang, 2014). The qualitative data suggest that the ACs were better than the PPTs in boosting learners' motivation, which is as significant as test scores. Moreover, the former also seemed to have added variety and fun to the lessons and broke the routine in the class.

The ACs also transformed the way the participants viewed punctuation in writing. This is perhaps more important than learning how to use punctuation marks because correct use of punctuation is sometimes a matter of awareness rather than knowledge. In this respect, one of the most conspicuous results to emerge from the data was that the ACs helped raise awareness of punctuation among the participants. A possible reason for this could be the presentation of the rules in a simulated environment. Moreover, as it was the first time the participants were taught using a combination of digital tools, it seemed a novel experience for them.

In short, as put forward by Johnson (2006), the evidence found in the present study confirms that learners love watching ACs. This seemed so, although the ACs in the current study were created by the teacher rather than professionals, and this is good news for teachers aiming to put their creativity into practice. However, as learners are exposed to higher quality digital learning materials and tools, they might find it less engaging to work with such materials. This implies that overuse of such materials might undermine their instructional value.

As there appears to be a thin line between entertainment and study with respect to watching the ACs, exercises and partial forum discussions helped the participants in the treatment group stay



awake rather than fall prey to the tendency to watch videos passively as cautioned by Walker (1999, as cited in Król, 2013). However, as with Su and Liang's (2014) study, one participant raised her voice against watching videos following a lesson format characterised by frequent pauses, repetitive watching, and exercises. Moreover, the language used in the ACs and the pace of the speech suited the level of the participants. Unlike the participants in Su and Liang's (2014) study, who had difficulty understanding the language of the cartoons, no participants in the present study reported problems regarding comprehensibility or intelligibility of the text-to-speech voices. This might imply that, despite their downsides, they could be used to produce multimedia materials in language instruction, particularly when there are no native speakers.

Finally, it is also worth focusing on the statistically significant improvement in the control group from the pretest to posttest and positive perceptions of the learning design and materials. As noted earlier, the only difference between the two learning conditions was the main delivery material; everything else was identical. Although the PPTs did not grab the learners' attention (as they are a mainstream instructional tool in Turkey), an instructional design presented on a learning management system supported with exercises seemed interesting because they had never studied English by using such innovative and controlled online activities. This also explains why the digitally less experienced interview participants viewed the learning experience more positively (see Figure 3). Therefore, one could hypothesize that if they had watched the PPTs as a whole class and done paper-based exercises, they might have viewed the learning experience much less positively.

## 6. Conclusions

The present study attempts to add a new item to the agenda of user-friendly technologies in EFL. A significant upshot of this study is that ACs could help EFL learners increase their knowledge of punctuation rules and raise their awareness of correct punctuation use. This is particularly important, as awareness is a significant asset for applying rules of punctuation. The evidence from this study suggests that they could be used not only to motivate students to study, but also to transform their views about a particular subject matter. Taken together, the results of this study suggest that ACs with follow-up activities might add variety into instruction and function as a shortcut to explaining topics of relative difficulty.

### 6.1 Pedagogical implications

As the results of the current study suggested, a viable option for developing an awareness of punctuation might be the use of ACs to draw students' attention, but they should be as short as possible, particularly when they introduce lots of information. Also, accompanying exercises are strongly recommended to prevent learners from seeing cartoons as a tool for mere entertainment. However, a downside of creating tailor-made ACs might be the excessive amount of time they necessitate. Therefore, teachers could cooperate with colleagues to create ACs for commonly experienced problems. Despite such difficulties involved in creating tailor-made ACs, the innovation and variety seem worth the time and effort spared for them. Therefore, they might also be used to teach other challenging issues in L2 pedagogy in line with the cultural and linguistic background of learners.

### 6.2 Limitations and recommendations for further research

There are several limitations of the study. First, teacher perceptions could have enriched discussion and helped paint a more complete picture. Second, although the researcher collected survey data from both groups, the interviews were carried out only with participants from the experimental group. Interview data from the control group could have helped understand the quantitative differences between the experimental and control group better. In addition, a second control group working with only paper-based activities could have provided deeper comparative data. Finally, as the researcher was the teacher of some of the interview participants, they might

have tried to please him, although he tried hard to establish rapport during the interviews and stressed that he needed objective responses.

Prospective researchers could ask their students to create their own ACs and investigate how student-generated cartoons could fit in L2 pedagogy. Further research could also focus on such issues related with text-to-speech technologies as intelligibility, speech rate, accent, and so forth. Finally, prospective researchers could also extend the use of ACs to include their use in teaching different language skills/areas.

**Supplementary materials.** For supplementary materials referred to in this article, please visit <https://doi.org/10.1017/S0958344018000046>

**Acknowledgment.** The researcher gratefully acknowledges the contributions of anonymous reviewers who helped improve the quality of this paper.

**Ethical statement.** The author declares no conflicts of interest that might affect the results and their interpretation in the study. He also states that ethical principles of conducting scholarly research were taken into account, during the process of data collection, data analysis, and reporting.

## References

- Alaba, S. O. (2014) Innovative method of teaching English language in primary schools using adaptive interactive animated cartoons. *Journal of Educational and Social Research*, 4(6): 75–82. <https://doi.org/10.5901/jesr.2014.v4n6p75>
- Alamin, A. & Ahmed, S. (2012) Syntactical and punctuation errors: An analysis of technical writing of university students science college, Taif University, KSA. *English Language Teaching*, 5(5): 2–8. <https://doi.org/10.5539/elt.v5n5p2>
- Altunay, D. (2009) *Use of connectives in written discourse: A study at an ELT department in Turkey*. Anadolu University, Eskişehir, unpublished PhD.
- Arikan, A. & Taraf, H. U. (2010) Contextualizing young learners' English lessons with cartoons: Focus on grammar and vocabulary. *Procedia - Social and Behavioral Sciences*, 2(2): 5212–5215. <https://doi.org/10.1016/j.sbspro.2010.03.848>
- Backscheider, P. (1972) Punctuation for the reader: A teaching approach. *The English Journal*, 61(6): 874–877. <https://doi.org/10.2307/813995>
- Benzer, A. (2010) Prospective teachers' proficiency in punctuation rules and opinions related to punctuation problems. *Procedia - Social and Behavioral Sciences*, 2(2): 1878–1883. <https://doi.org/10.1016/j.sbspro.2010.03.1002>
- Boucheix, J.-M. & Guignard, H. (2005) What animated illustrations conditions can improve technical document comprehension in young students? Format, signaling and control of the presentation. *European Journal of Psychology of Education*, 20(4): 369–388. <https://doi.org/10.1007/BF03173563>
- Chandler, P. & Sweller, J. (1991) Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4): 293–332. [https://doi.org/10.1207/s1532690xci0804\\_2](https://doi.org/10.1207/s1532690xci0804_2)
- Coffin, T. P. (1951) An aid to the teaching of punctuation. *College English*, 12(4): 216–219. <https://doi.org/10.2307/372628>
- Cohen, J. (1988) *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Creswell, J. W. (2012) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.
- Creswell, J. W. & Plano Clark, V. L. (2011) *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.
- Dawkins, J. (2003) Teaching meaning-based punctuation. *Teaching English in the Two-Year College*, 31(2): 154–162.
- DeCoursey, C. A. (2012) Trialing cartoons: Teachers' attitudes towards animation as an ELT instructional tool. *Computers & Education*, 59(2): 436–448. <https://doi.org/10.1016/j.compedu.2011.09.005>
- Elkılıç, G., Han, T. & Aydın, S. (2009) Punctuation and capitalization errors of Turkish EFL students in composition classes: An evidence of L1 interference. *First International Symposium on Sustainable Development, Sarajevo, Bosnia and Herzegovina*. International Burch University, 9–10 June.
- Gauthier, L. R. (1993) A strategy to increase punctuation awareness. *Journal of Reading*, 36(5): 401–402.
- Given, L. M. (Ed.). (2008) *The Sage encyclopaedia of qualitative research methods*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412963909>
- Hirvela, A., Nussbaum, A. & Pierson, H. (2012) ESL students' attitudes toward punctuation. *System*, 40(1): 11–23. <https://doi.org/10.1016/j.system.2012.01.006>
- Johnson, D. (2006) *Using video cartoons to encourage EFL practice*. <https://donjohnson.wordpress.com/academic-writing/using-video-cartoons-to-encourage-efl-practice/>
- Karakaş, A. & Sarıçoban, A. (2012) The impact of watching subtitled animated cartoons on incidental vocabulary learning of ELT students. *Teaching English with Technology*, 12(4): 3–15.

- Khodashenas, M. R., Farahani, S. K. & Alishahi, Z. (2014) Flash cards versus animated cartoons: A comparative study in vocabulary teaching and learning. *International Journal of Language Learning and Applied Linguistics World*, 5(4): 290–298.
- Kirkgöz, Y. (2010) An analysis of written errors of Turkish adult learners of English. *Procedia - Social and Behavioral Sciences*, 2(2): 4352–4358. <https://doi.org/10.1016/j.sbspro.2010.03.692>
- Kristiansen, A. (2001) *Animation and teaching: Enhancing subjects from the curriculum by using "The Simpsons" in high school English teaching*. <https://www.simpsonsarchive.com/other/papers/ak.paper.html>
- Król, W. (2013) Research communiqué on the use of animated cartoons in teaching English to children with disorders and disabilities. *Procedia - Social and Behavioral Sciences*, 106: 2259–2268. <https://doi.org/10.1016/j.sbspro.2013.12.257>
- Lee, W. (2009) Making English lessons engaging through video materials supported with advance organizers and prediction activities. *TESL Reporter*, 42(2): 57–74.
- Malone, T. W. (1981) Toward a theory of intrinsically motivating instruction. *Cognitive Science*, 5(4): 333–369. [https://doi.org/10.1207/s15516709cog0504\\_2](https://doi.org/10.1207/s15516709cog0504_2)
- Mayer, R. E. & Anderson, R. B. (1992) The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84(4): 444–452. <https://doi.org/10.1037/0022-0663.84.4.444>
- Mayer, R. E. & Moreno, R. (2002) Animation as an aid to multimedia learning. *Educational Psychology Review*, 14(1): 87–99. <https://doi.org/10.1023/A:1013184611077>
- Mayer, R. E. & Sims, V. K. (1994) For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86(3): 389–401. <https://doi.org/10.1037/0022-0663.86.3.389>
- Meilleur, R. (2004) The Simpsons in Japan: A lesson on stereotypes. *Internet TESL Journal*, 10(3). <http://iteslj.org/Lessons/Meilleur-Simpsons.html>
- Moulic, M. (2012) *Animation film to develop effective listening skills in the classroom*. <https://doi.org/10.7718/iamure.ije.v2i1.166>
- Paivio, A. (1991) Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, 45(3): 255–287. <https://doi.org/10.1037/h0084295>
- Patton, M. Q. (2002) *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks: Sage.
- Perry, FL, Jr. (2005) *Research in applied linguistics: Becoming a discerning consumer*. Mahwah: Lawrence Erlbaum.
- Porte, G. K. (2002) *Appraising research in second language learning: A practical approach to critical analysis of quantitative research*. Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.3>
- Schnotz, W., Böckheler, J. & Grzondziel, H. (1999) Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education*, 14(2): 245–265. <https://doi.org/10.1007/BF03172968>
- Stratton, M. T. & Julien, M. (2014) Xtranormal learning for millennials: An innovative tool for group projects. *Journal of Management Education*, 38(2): 259–281. <https://doi.org/10.1177/1052562913504923>
- Strauss, A. L. & Corbin, J. (1998) *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). London: Sage.
- Su, S.-C. & Liang, E. (2014) Elementary students' motivation and attitudes toward English animated cartoons at a cram school. *International Journal of Arts and Commerce*, 3(5): 15–36.
- Su, S.-C. & Liang, E. (2015) Using subtitled animated cartoons and textbook-based CDs to test elementary students' English listening and reading comprehension in a cram school. *International Journal of Applied Linguistics & English Literature*, 4(1): 106–114.
- Tracy, S. J. (2013) *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. Hoboken: Wiley-Blackwell.
- Tversky, B., Morrison, J. B. & Betrancourt, M. (2002) Animation: Can it facilitate? *International Journals of Human-Computer Studies*, 57(4): 247–262. <https://doi.org/10.1006/ijhc.2002.1017>
- Waugh, D. (1998) Practical approaches to teaching punctuation in the primary school. *Reading*, 32(2): 14–17. <http://dx.doi.org/10.1111/1467-9345.00083>

### About the author

Dr Arif Bakla is a lecturer at the Department of Foreign Language Education, Cumhuriyet University, Turkey. He holds a PhD degree in ELT. Among his research interests are web-based language instruction, LMSs, authoringware, and feedback in L2 writing. He is the author of the book *Putting Pen to Paper: Academic Paragraph Writing*.

Author ORCID.  Arif Bakla, <http://orcid.org/0000-0001-5412-4330>