



# On the adequacy of the experimental approach to construct validation: the case of advertising literacy



Mohammad Salehi<sup>\*</sup>, Afsaneh Farhang

*Languages and Linguistics Center, Sharif University of Technology, Tehran, Iran*

## ARTICLE INFO

**Keyword:**  
Education

## ABSTRACT

Given the significant number of the studies conducted to explore construct validity, adopting various approaches, and the paucity of the studies that employed the experimental approach to validation, the current study was aimed at exploring the adequacy of the experimental approach to construct validation, drawing upon advertising literacy as the construct of the researchers' concern. To this end, the researchers pursued the experimentation, test administration, and validation phases, respectively. To identify the effect of the experimentation, an independent samples t-test was run; there was a significant difference between the experimental and control groups. The closing phase was devoted to verification of the construct validity of the test. Hence, it was referred to the divergence of scores on the test of advertising literacy and a midterm examination of general English proficiency. The results indicated a low correlation between the two sets of scores. Therefore, it could be argued for the desirable degree of construct validity. Accordingly, the adequacy of the approach is evidentially verified.

## 1. Introduction

### 1.1. Overview

During the decades, with the recognition of testing as an indispensable part of teaching and learning program, a large body of research has been recorded on the investigation of validity as an essential quality of the tests. Although numerous validation studies have been conducted using various approaches, such as Anderson et al. (1991); Bachman and Palmer (1982); Rezaee and Salehi (2008); Cohen and Upton (2007), there are few studies that employed the experimental approach to construct validation. The current study is mainly aimed at development and validation of a measure to tap advertising literacy adopting the experimental approach to construct validation. As the title suggests, the central goal of this study is to determine whether the experimental approach may provide adequate pieces of evidence for verification or falsification of the assertion that a test, the researchers developed to measure a theoretical construct, enjoys a fair degree of construct validity. Therefore, it could be argued that this study is primarily intended to explore the adequacy of the experimental approach to validation with a focus on advertising literacy, as the construct of the researchers' concern. Thus, the notion of advertising literacy serves as a tool whose measurement would assist the researchers in exploring the adequacy of the

experimental approach that the researchers employed.

In the experimental approach to construct validation, according to Bachman (1990), the researcher pursues three sequential phases including experimentation, test administration, and validation, respectively. Based on the effect of the experimentation phase, being already known to the researcher, the unknown validity of the test would be accordingly explored. Regarding the already known effect of the experimentation, the researchers refer to an abundance of evidence from research (Roschelle et al., 2005; Chinnery, 2006; Wardhaugh, 2003; and Stamatelou, 2015) supporting the assumption that getting the learners familiarized with authentic advertisements, which is produced by native speakers in a native speaker community (Porter and Roberts, 1981), may assist their assimilation of foreign language advertisements. In other words, the extent to which the learners may correctly identify the advertising message behind the language specifically used in advertising discourse, is influenced by their former exposure to advertising content. The validation phase is conducted, using independent samples t-test as well as referring to the notion of divergent validity to ensure that the test precisely measures the construct of the researchers' concern. The answer to the research question concerning the adequacy of the experimental approach is highly dependent upon the extent to which the researchers could verify or falsify the construct validity of the test primarily designed to tap advertising literacy on the basis of the experimental approach to

<sup>\*</sup> Corresponding author.

E-mail address: [m\\_salehi@sharif.edu](mailto:m_salehi@sharif.edu) (M. Salehi).

validation.

## 2. Background

### 2.1. Theoretical background

#### 2.1.1. Construct validity and approaches to validation

As it was previously pointed out, a measure of a psychological construct may possess high degrees of construct validity if the scores on the test reflect hypothetical characteristics and qualities which are constructed to account for the observable performance on the test. According to [Ary et al. \(2013\)](#), test developers have to provide adequate evidence to ensure that the scores really reflect the construct in question. They suggested one should begin with a definition of the construct to be measured in the test and collect any logical and empirical evidence “that supports the assertion that a test measures the construct as defined and nothing else” ([Ary et al., 2013](#), p. 248). They further recommend some strategies to gather construct related evidence, as follows: 1. Related measures studies, 2. Known-groups technique, 3. Intervention studies, 4. Internal structure studies, 5. Studies of response processes.

According to [Bachman \(1990\)](#), as far as the abilities to be measured are not directly observable, they can be inferred based on the performance observable performance. He regarded these abilities as theoretical, in that in a hypothetical condition one could say that these abilities affect language use as well as the performance on language tests ([Bachman, 1990](#)). Given the significance of the extent to which one may make true inferences based on test performance, he added, in construct validation, therefore, “we seek to provide evidence that supports specific inferences about relationships between constructs and test scores; that is, in conducting construct validation, we are empirically testing hypothesized relationships between test scores and abilities” ([Bachman, 1990](#), p. 256).

In order for the current study to be supported by a strong theoretical background, one should refer to the almost closely relevant theoretical argument by [Bachman \(1990\)](#) concerning the experimental approach to construct validation; however, it is essential to refer to the variety of approaches to construct validation procedures before beginning the discussion regarding the researchers’ preferred approach.

As [Messick \(1989\)](#) argued, the test developer may probably collect several types of empirical evidence to support construct validity of the tests. These might include all or any of the following evidence: (1) to examine the correlational patterns among test scores and item scores, and between characteristics of items and tests and scores on items and tests; (2) to analyze and to model the processes underlying test performance; (3) to study the differences between groups; (4) to study the changes over time; (5) to investigate the effects of experimental treatment ([Messick, 1989](#)). [Bachman \(1990\)](#), however, proposed two other types of powerful evidence, namely, correlational and experimental. He further introduced factor analysis as one of the commonly used procedures for interpreting a large number of correlations.

Also, [Campbell and Fiske \(1959\)](#) described the classic approach to design correlational studies on construct validation, that is to say, the multitrait multimethod (MTMM) matrix. According to this approach, the tests are included in the matrix to combine multiple traits with multiple methods as each considers each measure is considered as a combination of trait and method ([Campbell and Fiske, 1959](#)). “One advantage of the MTMM design is that it permits the investigator to examine patterns of both convergence and discrimination among correlations” ([Bachman, 1990](#), p. 263).

[Messick \(1989\)](#) also underlined the importance of investigating the processes involved in test taking performance, stating that: “in numerous applications of techniques for studying the process, it became clear that different individuals performed the same task in different ways and that even the same individual might perform in a different manner across items or on different occasions. . . . That is, individuals differ consistently in their strategies and styles of task performance [emphasis in original]”

([Messick, 1989](#), p. 54).

[Messick \(1989\)](#) further discussed a number of approaches to analyzing the test taking process, as follows: protocol analysis and computer modeling, the analysis of response times and mathematical modeling of tests, the analysis of reasons given by test takers for choosing a particular answer, and the analysis of systematic errors ([Messick, 1989](#)).

As the current study has been conducted in line with the core principles of the experimental approach to construct validation, it is worthwhile to note the distinction between experimental studies and construct validation on the basis of experimentation. According to [Bachman \(1990\)](#), the former is aimed at exploring the potential effects of an experimental treatment in which the researcher employs at least one treatment whose effects are unknown, while evaluating the subsequent effects of the treatment in terms of one’s performance on a valid indicator whose effect is already known; such as a test of language ability. However, as [Bachman \(1990\)](#) argued, in the latter “it is the test, and not the treatment, that is of primary interest. Therefore, since we are examining a test whose validity is as yet unknown, we must employ a treatment whose effects are known” ([Bachman, 1990](#), p. 267).

In sum, one should acknowledge that obtaining empirical evidence for construct validity of a test is a demanding task for test developers. Naturally, such a cumbersome process prerequisites continuous accumulation of data from a variety of sources and as [Messick \(1989\)](#) contends, different types of validity, such as content and concurrent, should be viewed only as means to ensure construct validity.

#### 2.1.2. Advertisement instruction, advertising literacy, and L2 development

[Malmelin \(2010\)](#) regarded advertising as a unique form of communication whose interpretation requires specific literacy skills that largely differ from those needed in the assimilation and production of other symbolic forms of communication. According to [Malmelin \(2010\)](#), one approach to the illustration of advertising literacy, as a construct, is by means of modeling its various dimensions. Hence, as he noted, advertising literacy can be divided into four dimensions of informational literacy, aesthetic literacy, rhetorical, and promotional literacy. Also, “advertising literacy can be approached from three different vantage-points; from the point of view of media education and research, from the consumer point of view, and finally, from the communicator’s perspective” ([Malmelin, 2010](#), p. 139).

[Rozendaal et al. \(2011\)](#) highlighted the need to extend the one-dimensional conceptualization of advertising literacy as the ability to make sense of the message behind advertising content, and proposed two extra dimensions of advertising literacy as follows: attitudinal advertising literacy and advertising literacy performance taking into account the actual use of conceptual advertising literacy knowledge when confronting with advertisements.

Moreover, [Gill \(2013\)](#) referred to the teaching function of advertising asserting that advertising is akin to teaching. As the teacher, in the first place, attempts to steer the learners’ attention and interest towards the subject matter and secondly prescribes the means by which the relevant figures, facts, and notions could be stored in memory for later recall. Furthermore, he considered a similarity between a teacher and an advertiser who calls attention to his message regarding the product’s advantages using different strategies to influence his audience. He asserted that repetition plays a significant role in enhancing the assimilation of various points about the product. Likewise, a teacher makes use of a variety of techniques to make his message facily received, stored and retrieved ([Gill, 2013](#)).

Also, the notion of advertising literacy and its potential effect on specific age groups were discussed by another research group. [Hudders et al. \(2017\)](#) explored the issues that arise in children’s encounter with embedded advertising and underscored the need to activate and apply their advertising literacy skills in a way that they deal with advertisements in a conscious and critical manner. They further unraveled various dimensions of advertising literacy in the context of embedded advertising

when commercial content is embedded into the media content. They provided an in depth theoretical framework for investigating how children process embedded advertisements. And finally, in their conceptual paper some research directions concerning how to expand children's dispositional and situational advertising literacy skills were discussed, so as to strengthen the way they encounter embedded advertisements (Huders et al., 2017).

## 2.2. Empirical studies

### 2.2.1. Construct validity and validation studies

Anderson et al. (1991) conducted an in-depth exploratory study into the construct validity of a reading comprehension test. Their study served as a methodological exploration in the use of information elicited from think-aloud protocols and more commonly used types of information on test content and test performance in the investigation of construct validity.

In order to provide empirical support for a model that incorporates all the components of communicative competence, Bachman and Palmer (1982) examined the construct validity of some tests of components of communicative competence including linguistic competence, pragmatic competence, and sociolinguistic competence using a multitrait-multimethod design. For the investigation of construct validity of tests to tap each of the hypothesized traits, they used four methods: an oral interview, a writing sample, a multiple-choice test, and a self-rating. They also applied confirmatory factor analysis to ensure the plausibility of several causal models, involving from one to three trait factors. The findings of their study produced evidence for a substantial amount of general factors influencing all the measures in the study (Bachman and Palmer, 1982).

One of the most related studies conducted within the realm of construct validation is by Rezaee and Salehi (2008) investigating the construct validity of a test used for admission purposes. In order to determine whether the test enjoys an acceptable degree of construct validity, the design of multitrait-multimethod as a prime tool was used. In their study, the two traits used were grammar and vocabulary and the two methods employed were multiple choice and contextualization. The researchers also referred to the notions of convergent and divergent/discriminant validity which are logically related to the convergence of scores obtained through different methods that measure the same trait and divergence of scores obtained through the same methods that measure different traits, respectively. The researchers also conducted two monotrait-multimethod analysis; one was designed to test grammar through more than one method and the other to test vocabulary using more than one method. The findings of the study revealed that the test possessed both convergent and divergent validities. In other words, "it was shown that the same trait measured through the same methods had a higher correlation than that of the same traits having been tested through different methods" (Rezaee and Salehi, 2008, p. 108).

As Bachman (1990) asserted, despite the fact that there are numerous studies adopting correlational approaches to investigate construct validation, there have been very few construct validation studies of language tests that employ an experimental approach as a framework. However, some of the few studies that employed the experimental approach to validation, including Palmer (1979), Fischer (1981), and Chacevych et al. (1982), are discussed in the current section.

One study, which is not directly conducted as a construct validation study, as Bachman (1990) noted, but which lends itself to such interpretation was performed by Palmer (1979) providing data from three studies to support the hypothesis regarding the presence of two distinctive types of language control: compartmentalized and integrated control. One of the three studies was a study of individual differences; he examined the relationship between linguistic competence and communication ability for two groups of second language learners. And the other two studies were aimed at controlling the experiments in foreign language instruction. The findings supported the fact that differences

between 'discrete-point accuracy drills' and 'communication activities as instructional treatments contribute to differences in performance on different types of language tests of achievement and communication.

The results of another study within the experimental approach's framework were reported by Fischer (1981), suggesting that one's better performance on a communicative competence component of a final examination may be achieved through actual practice in taking what he calls communicative tests.

Another relevant Example can be the study Chacevych et al. (1982) carried out. He found that observed differences in performance on ESL cloze tests follow from prior exposure to passage content in the native language. To put it in another way, the diversity of scores achieved on the ESL cloze tests was significantly related to the experimentation phase of the study during which the learners received the passage content in the native language, thus raising questions about the constructs measured by the cloze. "While the number of such studies is quite small, they are illustrative of the application of the experimental approach to construct validation" (Bachman, 1990, p. 268).

### 2.2.2. Advertising language and L2 development

Hobbs et al. (2014) supported the development of the learners' critical thinking and English language skills, thorough advertising analysis activities using key critical questions for analyzing a variety of media messages and documents designed to strengthen students' vocabulary, reading, and discussion skills to build inferential thinking, and critical analysis skills. In their study, the activities provided a meaningful opportunity for students to practice their all four language skills in English while supporting the development of abstract thinking, balancing the learning of new academic language and cultural knowledge and connecting with students' prior knowledge, home cultures and everyday experience (Hobbs et al., 2014).

Belova et al. (2016), in their research which was aimed at focusing on the roles for advertisements, which differ widely in their complexity and contributions to the media literacy in science education and encompass the two major media literacy goals of reading and writing media, they identified the significance of advertisements in fostering the sense of motivation, provoking science-related questions, which often result in experiments, encouraging reflection on science and science-related factual content (Belova et al., 2016).

Having centered upon the process of language assimilation, Gass (1997), argued "learners who have already been taught discourse features at discourse level using whole written advertisements may compare their L1 and L2; comparison of L1 and L2 leads to intake the processing of assimilating linguistic material" (Gass, 1997, pp. 3–5).

When it comes to the measurement of advertising literacy as construct, the notion of advertising literacy should be explicated in the first place. Rozendaal et al. (2016) attempted to develop a reliable and valid instrument to measure children's advertising literacy. Accordingly, the existing conceptualizations of advertising literacy primarily center upon conceptual knowledge of advertising which is basically defined as the ability to understand and recognize the message behind the advertisements. The multi-dimensional conceptualization of advertising literacy, is the basis for the development of such a scale to measure advertising literacy. Despite the equal importance of all three dimensions including conceptual, attitudinal, and performance advertising literacy, Rozendaal et al. (2016) mainly focused on the measurement of conceptual and attitudinal dimensions of children's advertising literacy only. Therefore, they attempted to measure the ability to retrieve and apply advertising knowledge and attitudes. They tried to develop an advertising literacy skill scale targeting children between 8 to 12 as it was generally assumed that in this age span, children go through some significant developmental changes in terms of socio-cognitive and information processing capabilities which probably exert considerable effects on one's advertising literacy development. To determine whether the test of the advertising literacy components indeed measure the two dimensions of conceptual and attitudinal advertising literacy, the

researchers created a 39-item version of the ALS-c.

They also checked the validity and reliability of the scale to test its assumed structure and to ensure the precision in measuring respondents' scores.

Finally, they could develop an almost reliable and valid scale to tap children's advertising literacy. They argued that "such a measurement instrument is needed to investigate the general development of advertising literacy in children and to study the role of advertising literacy in children's susceptibility to advertising" (Rozenaal et al., 2016, p. 21).

As the aims and findings of the above mentioned studies suggest, the concept of advertising literacy and construct validation of the test designed to measure this concept are of central importance to fall within the ambit of the current study. Thus, the research questions to be addressed in this area are as follows:

- 1) Does the inclusion of authentic advertisements in both print and video versions make a significant difference in the learners' performance on the test of advertising literacy?
- 2) Does the researcher developed test of advertising literacy enjoy a desirable degree of construct validity?
- 3) Does the Experimental Approach to construct validation provide adequate pieces of evidence for verification or falsification of the assumed validity of the tests?

### 3. Methodology

#### 3.1. Participants

In this study, the researchers selected eight groups of learners studying general English as a foreign language at Sharif University of Technology. Each group comprised twenty B.A. students of various fields, both males and females of different ages ranging from 20 to 25. Prior to data collection procedure, informal consent was obtained from the participants. As will be illustrated in more details, one of the groups served as the experimental group to which the treatment, namely authentic advertisement instruction, was given. The other seven groups served as the control groups that received the usual instructional program for learning English as a course. Moreover, as it was attempted to select the group members in a way that the participants be in approximately similar levels of language proficiency; relying upon the learners' scores on a midterm exam, which had been identified as virtually intermediate in level, all the participants were considered as mostly intermediate in terms of general English proficiency. Therefore, the materials used during the experiment would supposedly fit the learners' level of proficiency. The ethical committee named "Sharif Language Center Ethical Committee" approved the study prior to the experimentation.

#### 3.2. Stimulus materials

The materials used during the experiment comprised a selection of authentic English advertisements including both print and video versions. The print adverts were predominantly taken from Reader's Digest book as the main source of the advertisements containing a large body of vocabulary items and expressions specifically used in advertising discourse. In addition, some other parts of print adverts used as a source for both instruction and test construction in the subsequent stages, were as follows: New York Times, Financial Times, Bloomberg BusinessWeek in which a large number of advertising sections are covered, introducing a variety of products and services.

The video versions of advertisements were mainly chosen from online platforms on which a significant number of authentic adverts in a variety of accents of English are uploaded on a daily basis and in which numerous products or services are introduced to the world. However, it could be said that near ninety percent of the videos used in the experiment were downloaded from among those uploaded on [www.youtube.com](http://www.youtube.com) as one of the most popular online platforms on which approximately

one hour of video is uploaded every second. The videos were virtually similar in terms of length; all of them took near five to eight minutes. Regarding the content, it was initially attempted to cover an almost various range of topics. However, as the outcomes of the piloting stage indicated that the majority of the learners showed significant interest in videos whose content was related to technology and technological advances, such as the adverts which introduced a new breed of mobile phones or a car with outstanding, the materials were modified in terms of content, length, and nature according to the young generation's interests and needs.

#### 3.3. Data collection procedure

The major aim of the researchers, as the title suggests, was to delve into the process through which the assumed construct validity of the tests is verified or falsified. Therefore, it could be decided whether the experimental approach may adequately provide sufficient pieces of evidence to investigate the notion of construct validity. As for the multi-phasic nature of the experimental approach to validation, the researchers began with the initial phase of the study, namely, the experimentation phase, whose results might serve as a baseline for the other stages to be subsequently followed.

To determine whether the inclusion of advertisements in the language learning syllabus exert any effect on the learners' future performance on the test designed to tap advertising literacy, the opening phase of the study was conducted to let the learners have a closer look at English advertisements in which myriads of words and phrases specific to advertising discourse were used. What is known, as the central premise of the experimental approach, is that the learners who receive advertising instruction may probably outperform their peers in terms of both knowledge of advertising items and the degree to which they find the advertising language almost tangible, compared to those with no previous familiarity with such a discourse. However, to explore the potential effect of the experimentation phase, the following phase of the study requires the administration of the researcher developed test of advertising literacy designed to measure the construct of the researchers' concern.

It is noteworthy to stress the fact that validation process is an indispensable part of each test development procedure; thus, the researchers had to focus on the learners' scores obtained through the test to decide if the newly developed test of advertising literacy truly indicates the difference between the experimental and control groups in terms of their performance on the test. As it was mentioned, it is expected that the learners being familiar with advertising language may probably achieve higher scores compared with others. Therefore, if the learners' scores on the test supported the same assumption, one could argue that the test enjoys a desirable degree of construct validity; however, if the learners' scores did not highlight the existing difference between the experimental group and the control group in terms of advertising literacy, one may strongly argue that the test lacks the adequate degree of validity. Thus, the following phase of the research is primarily concerned with construct validation of the test of advertising literacy.

##### 3.3.1. The experimentation (phase I)

As it was discussed earlier, the experimental stage was pursued during eight sessions of instruction taking around twenty minutes per session and near 160 minutes overall. The first group of learners, the experimental group, received the explicit treatment. During the experiment, all the learners were initially given some printed advertisements being authentic and containing words and expressions frequently used in advertising language. As to the researchers' explicit instruction, all the words and expressions to be learnt had been already underlined and explicitly defined in the class. Secondly, the advertising videos were played for the learners. In each session, only one video was played and each video was played twice to ensure that the learners had listened to

the expressions of the researchers' focus. In this stage, the learners were asked about the content of the video through posing a number of questions regarding the product or service having been advertised. The meaning of the unknown words and phrases were also clearly defined. The learners who received the experimental treatment tried to recognize the meaning of the unknown words according to the context, in other words, they were asked to clarify the message behind the advertising words and phrases by being even allowed to use their first language if they could not find the English equivalents to elaborate on the concepts.

### 3.3.2. Test administration (phase II)

The purpose of the second stage was to measure the underlying construct of advertising literacy. To meet this end, a test of advertising literacy was constructed (For the full version of the Test of Advertising Literacy, please see the 'Appendix'). The test comprised of twenty-eight items within the format of multiple-choice, paraphrasing, and true/false questions all of which were designed on the basis of authentic English advertisements introducing a variety of products and services. The test was designed to measure one's advertising literacy according to what the testees read and hear; to put it differently, the test items included a listening section based on the video versions of advertisements to investigate the learners' understanding of the advertising messages conveyed through auditory as well as visual means. Broadly similar to what was practiced during the experimentation phase, in this stage, the testees were asked to define what the advertiser meant by the underlined words, phrases, and sentences using either English or Persian, as the researchers' primary objective was concerned with the testees' ability to make sense of the advertising messages rather than the evaluation of their English proficiency level. It bears significance to mention that all the collected data were handled anonymously by assigning alphanumeric codes to each participant.

### 3.3.3. Validation (phase III)

Following the test administration phase, validation procedure began.

Concerning the validation process, which is to be discussed in more details, the learners' scores on a midterm examination of general English proficiency were also collected for the subsequent correlational procedures to make further reference to the notion of divergence of scores, as an indicator of construct validity.

## 3.4. Data analysis

In order to answer the research question regarding the potential changes that the inclusion of advertisements may effect in the learners' performance on the test of advertising literacy, the researchers initially ran an independent-samples t-test analysis to investigate the difference between the scores achieved by the twenty-member experimental group and the randomly selected control group from among the other seven groups having received no treatment, or advertising instruction.

The research question concerning the construct validity of the researcher developed test of advertising literacy, imposes the requirement of providing statistical evidence to argue for or against the validity of the newly developed test. To this end, the researchers referred to the notion of divergent/discriminant validity, assessing the significance of the correlation, if any, among the scores obtained from the test of advertising literacy and the midterm examination of general English proficiency. Presumably, if the correlation between the two sets of scores were approximately insignificant, it could be argued that the newly developed test enjoys a fair degree of construct validity since the two sets of tests measure no similar constructs. Moreover, the researcher developed test of advertising literacy was also placed under the scrutiny of a reliability analysis to ensure the consistency of scores achieved on the items. The results of the statistical analysis of the items will be discussed further in details.

## 4. Results

### 4.1. Results on the experimentation phase

The mean and standard deviation of scores on the 28-item test of advertising literacy administered for the experimental and control groups is presented in Table 1.

As presented in Table 2, an independent samples T-test was conducted to compare the scores on the test of advertising literacy, to gain insights into the potential influence of the treatment on the learners' subsequent performance on the test. As it has been illustrated in the table, there was a statistically significant difference in the scores for the experimental ( $M = 17.7$ ,  $SD = 1.8$ ) and control groups ( $M = 11.15$ ,  $SD = 2.3$ ;  $t(38) = 9.8$ ,  $p = .32$ , two-tailed).

### 4.2. Results on the validation phase (Divergent validity)

The researchers, during the validation phase, explored whether there existed a correlation between the scores achieved on the test of advertising literacy and on the midterm examination of general English proficiency as a course. Thus, the researchers referred to the notion of divergent/discriminant validity.

As Table 3 illustrates, two sets of scores achieved by 160 intermediate learners of English as a university course, were correlated. Given the notion of divergent/discriminant validity, one may observe that there is no substantial correlation among the two sets of scores, meaning that the newly developed test of advertising literacy, measures the construct of the researchers' interest and no wrong construct is being measured. The statistically low correlation among the scores provides divergent evidence to verify the construct validity of the test of advertising literacy.

### 4.3. Reliability analysis

The researchers also deemed it essential to check the reliability of the test of advertising literacy to determine the degree to which the items that make up this scale are all consistently measuring the same underlying attribute. A Cronbach's alpha approach was used to investigate the item variance reliability of the advertising literacy test.

According to Table 4, the scale has good internal consistency with a Cronbach alpha coefficient of .804. Values above .7 are considered acceptable; however, values above .8 are preferable (Pallant, 2011). In sum, as the table illustrates, the researcher made test of advertising literacy, enjoys a good coefficient of internal consistency.

### 4.4. Discussion

Since the early 1950s, "the concept of construct validity, the marked importance of validating the theoretical concept underlying a test, has been an important issue in educational and psychological measurement theory" (Noll and Scannell, 1972, p. 141). As it was illustrated in former sections, the principal objective behind this study was to gather sufficient pieces of evidence to support the adequacy of the experimental approach to explore the construct validity of the tests, having adopted the notion of advertising literacy as the construct of the researchers' concern;

**Table 1**

The mean and standard deviation of scores achieved on the test of advertising literacy by the experimental and control groups.

Group Statistics					
	Groups of Participants	N	Mean	Std. Deviation	Std. Error Mean
Advertising Literacy Scores	experimental	20	17.7000	1.89459	.42364
	control	20	11.1500	2.30046	.51440

As Table 1 illustrates, the experimental group has outperformed the control group.

**Table 2**  
The results on the Independent Samples T-Test analysis of the scores achieved by the two groups.

Independent Samples Test		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	Df	Sig. (2 tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Advertising Literacy Scores	Equal variances assumed	1.000	.324	9.829	38	.000	6.55000	.66639	5.20096	7.89904
	Equal variances not assumed			9.829	36.653	.000	6.55000	.66639	5.19933	7.90067

**Table 3**  
The correlation between the scores obtained through the test of Advertising Literacy and the midterm examination of general English proficiency as a course.

Correlations			
		Advertising Literacy Scores	Midterm Examination Scores
Advertising Literacy Scores	Pearson Correlation	1	.33 <sup>a</sup>
	Sig. (2-tailed)		.000
	N	160	160
Midterm Examination Scores	Pearson Correlation	.33 <sup>a</sup>	1
	Sig. (2-tailed)	.000	
	N	160	160

<sup>a</sup> Correlation is significant at the 0.01 level (2-tailed).

**Table 4**  
The reliability of the of the test of advertising literacy developed on the basis of the experimental approach to construct validation.

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.804	.803	28

interpreted another way, the notion of advertising literacy served as a tool to answer the question regarding the adequacy of the experimental approach to construct validation. The results obtained through the experimentation, test administration, and validation procedures, as specified in the experimental approach, turned out to statistically verify that relying on the experimental approach, one may accurately decide if the newly developed test possesses a fair degree of construct validity. As mentioned, the researchers provided a body of evidence to answer the three research questions. These pieces of evidence are discussed from various perspectives.

Regarding the effect of advertisement instruction at the experimentation phase, given the statistically significant difference in the scores for the experimental ( $M = 17.7, SD = 1.8$ ) and control groups ( $M = 11.15, SD = 2.3; t(38) = 9.8, p = .000$ , two-tailed), the potential influence of advertisement instruction in learners' performance on the subsequently administrated test is revealed. To explicate the reasons behind the experimental group's higher performance, it was referred to [Stamatelou, \(2015\)](#), who underscored the contribution of written advertisements to language learning. They could be used to teach macro-level and micro-level discourse features at the discourse level. Consequently, the learners can analyze authentic advertisements at discourse level in their classroom in their native language (L1) and in L2/English (target language) all over the world through various channels such as print and audiovisual advertisements ([Stamatelou, 2015](#)). A further justification for the difference between the means of scores for the two groups, would be the learners' overall familiarity with the advertising content and discourse during the experimentation phase, as the prior exposure to advertising content contributes to the preparedness of the testees to

expect some meaning beyond the literal sense of the words each advertisement may contain.

In response to the research question concerning the construct validity of the test, the observed divergence of scores on the two tests indicated that the researcher developed test successfully measured the trait of concern, almost distinct from what the midterm examination of general English proficiency was to assess. According to [Ary et al. \(2013\)](#) in the case of divergent evidence, the researcher attempts to show that no wrong construct is being measured ([Ary et al., 2013](#)). Almost in the same way, in a hypothetical condition, the high correlation between the two sets of scores would probably follow from the presence of some traits being shared between the two measures.

As argued earlier, importance has been attached to the concept of convergent and discriminant validation, as it has been to the concept of construct validation, even outside of language testing area. As [Palmer, Groot, and Trosper \(1979\)](#) in their justification of the influential role the concepts of convergent and discriminant validity, referred to [Campbell and Fiske \(1959\)](#) that brought together the concepts of validity having existed only separately into one conceptual framework. It was further referred to the most outstanding feature of the approach taken by [Campbell and Fiske](#) as their emphasis on discriminant validity:

“Discriminant validity refers to the notion that traits should be distinguishable from each other when measured by different methods. The situation is evidenced when the agreement between different measurement procedures for a trait is greater than the intercorrelation between that trait and others within the same measurement procedure” ([Palmer, Groot, & Trosper, 1979, p.58](#)).

The reason behind the low correlation coefficient between the two sets of scores, or to be specific, the observed divergence, could be explained by making a reference to the assumption that if a measure is valid, it should yield the same results when utilized across different methods, as [Campbell and Fiske \(1959\)](#) specified; if the results failed to converge, it would be suggested that the variations partially stem in part from the difference in measurement methods, therefore raising doubts about the validity of the measure. As [Table 3](#) illustrates, there is an approximately low correlation  $r = 0.3$  between the scores obtained through the two tests, indicating that the constructs of each measure's focus are not shared by the other measure and vice versa. Based on the observed divergence of scores the researchers may almost strongly argue, on the one hand, for the adequacy of the experimental approach based on which the three phases of experimentation, test administration, and validation were respectively pursued, and for the high applicability of the researcher developed test of advertising literacy as a tool to measure the conceptual advertising literacy on the other hand.

The statistically significant difference in the scores for the experimental ( $M = 17.7, SD = 1.8$ ) and control groups ( $M = 11.15, SD = 2.3; t(38) = 9.8, p = .32$ , two-tailed) can be traced back to the learners' prior familiarity with the advertising discourse and content. In the same way, [Hobbs et al., 2014](#) supported the development of the learners' critical thinking and English language skills, thorough advertising analysis activities using key critical questions for analyzing a variety of media messages and documents designed to strengthen the students'

vocabulary, reading, and discussion skills to build inferential thinking, and critical analysis skills. In their study, the activities provided a meaningful opportunity to practice all four language skills in English while supporting the development of abstract thinking, balancing the learning of new academic language and cultural knowledge and connecting with one's prior knowledge, home cultures, and everyday experience (Hobbs et al., 2014). What makes the current study virtually significant is the considerable attention paid to the learners' interests and needs in selection of advertising materials employed during the experimentation phase. At the opening steps of the study, the researchers attempted to cover a various range of topics in both print and audiovisual versions of advertisements; however, following the piloting stage according to the learners' reports, it came clear that the majority of the learners showed significant interest in videos whose content was related to technology and technological advances, as the adverts which introduced a new breed of mobile phone or a car with outstanding features that sounded attractive to young generation. Thus, as the researchers realized what the learners were most interested in, the materials were modified in terms of content and nature in line with their interests and needs. Consequently, as an outcome of the learners' interest and enthusiasm towards the videos or texts introducing a new product or service, the amount of their involvement and active participation in meaning-making process would be accelerated, and higher learning outcomes and better performance on the test would appear accordingly.

Regarding the authenticity of the instructional materials, according to the researchers' observations and the previously illustrated statistical procedures, adopting the print and audiovisual versions of authentic advertisements potentially enhanced the learners' motivation and interest throughout both phases of experimentation and test taking process in such a way that as the learners were informed of the authenticity of the advertisements, or in other words, as they realized that the advertisements introduced genuine products or services being accessible in real-life situation, and more significantly, as the native speaker produced pieces of language were provided, the amount of enthusiasm contributing to an active engagement with the meaning-making process, considerably maximized. Such finding, however, turned out to be in sharp contrast to that of Peacock (1997) who investigated the effect of the use of authentic materials on the learners' motivation. Results achieved out of a self-report questionnaire and two observation sheets indicated that while on-task behavior and observed motivation were significantly higher when authentic materials were used, the self-reported degree of motivation increased only over the last 12 of the 20 days of the research. Moreover, according to the results from self-reported questionnaires, the learners found authentic materials considerably less interesting than artificial materials (Peacock, 1997).

The final but crucial point to be mentioned is concerned with the researchers' informed choice of the conceptual dimension of advertising literacy which played a central role in the quality of the end-product of the study. The rationale behind the researchers' preference to specifically center upon the conceptual dimension of advertising literacy could be explained through the reference to the preferably high reliability of the test, with a Cronbach alpha coefficient of .804, following from the objectivity of the test items. Apparently, with the focus on attitudinal or performance advertising literacy, the test would inevitably contain some substantial degrees of subjectivity due to the lack of a clearly defined criterion for correctness.

## 5. Conclusion

### 5.1. Conclusion and suggestions for further research

This study was aimed at finding evidentially cogent answers to the research questions addressed earlier. Having adopted an experimental approach, the researchers attempted to answer the question regarding the influence that the instruction of authentic advertisements may exert on one's performance on the test of advertising literacy. Obviously, the

answer to the first research question fundamentally determined the subsequent phases of data collection to answer the following research questions.

The statistically collected evidence indicated that the experimental group having been exposed to the experimental treatment, achieved higher scores compared with the control group, in that, they had an approximately richer knowledge, or literacy, of the messages behind a variety of linguistic tools specifically used in the advertising discourse.

The answer to the second research question regarding the validity of the test of advertising literacy was highly contingent upon the correlational patterns among the scores achieved on the two sets of tests whose detailed description of administration procedures was discussed earlier in related sections. As the scores obtained through the test of advertising literacy diverged approximately highly from those on the midterm examination of general English proficiency, it was evidentially argued that the two sets of tests, measured distinct constructs. Moreover, the significantly low correlation among the scores on the test of advertising literacy administrated for the two experimental and control groups underlined the same notion in another way. The group having been exposed to advertising content performed highly better than the control group having received the normal instructional program, with no added advertising materials.

Given the fact that the answer to the last research question regarding the adequacy of the experimental approach to construct validation is inextricably linked to the previously answered research questions, it could be argued that the experimental approach to construct validation could adequately assist the researchers or future test designers in deciding if the newly developed test may differentiate between the experimental group and the control group who received no treatment that may potentially exert changes that lead to the enhancement or even detraction of the construct of interest.

The researchers' further reference to the notion of divergent/discriminant validity, also corroborated the adequacy of the experimental approach to investigate the construct validity of the test of advertising literacy. The approximately low correlation between the scores achieved on the midterm examination of general English proficiency and the test designed to measure the learners' advertising literacy highlighted the fact that the two sets of tests were fundamentally concerned with the measurement of two distinctively different constructs, and no similar traits were measured by the tests.

### 5.2. Further research

As pointed out in the earlier sections, while a plethora of studies has been conducted so as to investigate the construct validity of the newly developed tests, adopting various approaches to validation, there are almost few construct validation studies having adopted the experimental approach. The central concern of the current study was to decide if the test developers may rely on the experimental approach as an adequate means to ensure the validity of the tests. So far, despite a number of limitations and naturally existing complexities, the researchers could provide a body of evidence to support the assertion regarding the adequacy of the experimental approach to validation. However, Messick (1989) underlined the importance of investigating the processes involved in test taking performance. Therefore, a suggestion for the future research on the significance of construct validity would be supported by a virtually stronger argument if the approach taken to justify the presence or absence of the construct of interest provided adequate information on how test takers go through the process of test-taking performance. In other words, in this study the researchers managed to evidentially support the adequacy of the approach, however, it is essential to note that the statistically approved adequacy of the experimental approach is subject to the condition that the paramountcy of test-taking process, and test-taking strategies be not taken into account.

As discussed earlier, the notion of advertising literacy as the researchers' chief concern was viewed as a multidimensional concept.

Similarly, as according to Rozendaal et al. (2011), the need to extend the one-dimensional conceptualization of advertising literacy was underscored and three extra dimensions of advertising literacy were proposed, each comprising of several components: conceptual advertising literacy, attitudinal advertising, and advertising literacy performance (Rozendaal et al, 2011). Obviously, all the three dimensions including conceptual, attitudinal, and performance advertising literacy are equally important, though Rozendaal et al. (2016) mainly focused on the measurement of conceptual and attitudinal dimensions of children's advertising literacy. However, as the equal importance of the three dimensions of advertising literacy cannot be discarded, the future studies on the case of advertising literacy could be focused on the other dimensions on which there are very few, if any, theoretical or empirical arguments have been made.

## Declarations

### Author contribution statement

Mohammad Salehi, Afsaneh Farhang: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

### Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Competing interest statement

The authors declare no conflict of interest.

### Additional information

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2019.e01686>.

## References

Anderson, N.J., Bachman, L., Perkins, K., Cohen, A., 1991. An exploratory study into the construct validity of a reading comprehension test: triangulation of data sources. *Lang. Test.* 8 (1), 41–66.

Ary, D., Jacobs, L.C., Sorensen, C.K., Walker, D., 2013. *Introduction to Research in Education*. Cengage Learning, New York.

Bachman, L., 1990. *Fundamental Considerations in Language Testing*. Oxford University Press, New York.

Bachman, L.F., Palmer, A.S., 1982. The construct validation of some components of communicative proficiency. *Tesol Q.* 16 (4), 449–465.

Belova, N., Affeldt, F., Eilks, I., 2016. Using advertising as a teaching and learning medium in the science classroom. *Sch. Sci. Rev.* 97 (361), 86–92.

Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56 (2), 81–105.

Chacevych, A., Delva, C., Houle, F., Moon, R., Upshur, J.A., 1982. Language testing and content predictability. In: Paper presented at the Third Annual Language Testing Research Colloquium, Honolulu, Hawaii, May 1982.

Chinnery, G.M., 2006. Emerging technologies: going to the MALL (Mobile assisted language learning). *Lang. Learn. Technol.* 10 (1), 1–9.

Cohen, A.D., Upton, T.A., 2007. I want to go back to the text: response strategies on the reading subtest of the new TOEFL. *Lang. Test.* 24 (2), 209–250.

Fischer, R.A., 1981. Towards a practical model of a communicative competence testing program. In: Culhane, T. (Ed.), *English for Specific Purposes*, Vol 16. Modern English Publications, London, pp. 19–24.

Gass, S., 1997. *Input, Interaction and the Second Language Learner*. Lawrence Erlbaum, Mahwah, NJ.

Gill, L.E., 2013. *Advertising and Psychology*. Routledge, London.

Hobbs, R., He, H., Robbgrico, M., 2014. Seeing, believing, and learning to be skeptical: supporting language learning through advertising analysis activities. *TESOL J.* 6.

Hudders, L., De Pauw, P., Cauberghe, V., Panic, K., Zarouali, B., Rozendaal, E., 2017. Shedding new light on how advertising literacy can affect children's processing of embedded advertising formats: a future research agenda. *J. Advert.* 46 (2), 333–349.

Malmelin, N., 2010. What is advertising literacy? Exploring the dimensions of advertising literacy. *J. Vis. Lit.* 29 (2), 129–142.

Messick, S., 1989. Validity. In: Linn, R.L. (Ed.), *Educational Measurement*. Macmillan, New York, pp. 13–103.

Palmer, A.S., 1979. Compartmentalized and integrative control: an assessment of some evidence for two kinds of competence and implications for the classroom. *Lang. Learn.* 29 (1), 169–180.

Noll, V.H., Noll, R.P., Scannell, D.P., 1972. *Introductory readings in educational measurement*. Houghton Mifflin.

Pallant, J., 2011. *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using SPSS*. Allen & Unwin, Crows Nest, New South Wales.

Peacock, M., 1997. The effect of authentic materials on the motivation of EFL learners. *ELT J.* 51 (2), 144–156.

Porter, D., Roberts, J., 1981. Authentic listening activities. *ELT J.* 36 (1), 37–47.

Rezaee, A., Salehi, M., 2008. The construct validity of a language proficiency test: a multitrait-multimethod approach. *J. Teach. Engl. Lang.* 2 (8), 93–110.

Roschelle, J., Sharples, M., Chan, T.W., 2005. Introduction to the special issue on wireless and mobile technologies in education. *J. Comput. Assist. Learn.* 21 (1), 159–161.

Rozendaal, E., Lapierre, M.A., van Reijmersdal, E.A., Buijzen, M., 2011. Reconsidering advertising literacy as a defense against advertising effects. *Media Psychol.* 14 (1), 333–354.

Rozendaal, E., Oprea, S.J., Buijzen, M., 2016. Development and validation of a survey instrument to measure children's advertising literacy. *Media Psychol.* 19 (1), 72–100.

Stamatelou, M., 2015. A reflection on why and how the use of advertisements can enable language learning in the EFL classroom. *Open and Distance Learning* 8 (2), 78–88.

Wardhaugh, R., 2003. *An Introduction to Sociolinguistics*. Blackwell Publishing, Malden.