Effects of gender differences in the perceived learning quality of an AR-based learning scenario

Costin Pribeanu[1,2]

[1]National Institute for Research and Development in Informatics - ICI Bucharest

Bd. Maresal Averescu, nr.8-10, Bucharest, Romania

costin.pribeanu@ici.ro

[2]Academy of Romanian Scientists

ABSTRACT

In recent years there is an increasing interest in the quality of the educational software. Augmented Reality (AR) learning platforms have the potential to improve the understanding of specific concepts of a discipline. This paper aims to explore the effects of gender differences on the perceived learning quality of an AR-based learning scenario for Chemistry. A formatively measured model with causal indicators has been developed that has as focal construct the perceived Chemistry learning quality. An invariance analysis has been carried out to test if the model is invariant across. The results show that although the gender differences as regards the perceived learning quality are not significant, the boys and the girls have different perceptions as regards the contribution of the specific AR features to Chemistry learning.

Keywords

AR-based learning, causal indicators, invariance analysis, multidimensional model, educational quality.

INTRODUCTION

Augmented Reality (AR) learning tools have the potential to improve the understanding of specific concepts of a discipline and to enhance the motivation to learn by creating an attractive and enjoyable learning environment [6, 11, 14, 17].

In recent years there is an increasing interest to measure the quality of educational software. The perceived quality of an AR-based educational software could be seen as a general concept that manifests on specific dimensions such as the ergonomic, learning, and hedonic quality [14].

There are few studies that analyze the effect of the learner characteristics [6] on the learning outcomes of AR-based learning. The results of extant studies analyzing the gender differences in using computer-based learning systems are inconsistent as regards the differences and their statistical significance.

This paper aims to analyze the gender differences in the perceived learning quality of an AR-based learning scenario for Chemistry. A formatively measured model with causal indicators has been developed that has as focal construct the perceived learning quality. The causal indicators are the specific AR capabilities of the Chemistry learning scenario. The perceived learning quality is itself a bidimensional construct that manifests on the efficiency and usefulness of the learning scenario.

The model has been tested on the Augmented Reality Teaching Platform (ARTP) that has been developed in the framework of the ARiSE project [16].

RELATED WORK

AR-based learning

The AR-based educational systems are featuring typical capabilities, such as: 3D visualization, animation, vocal interface for learning and guidance, and haptic feedback. [4, 13, 17]. There are also specific AR features for a given learning scenario. For AR-based Chemistry learning, such capabilities may include augmentation of an atom structure, building a molecule from atoms, and simulation of chemical reactions [3, 4, 11].

AR environments are able to improve the understanding of spatial relationships through the visualization of abstract objects. Several studies mentioned the usefulness of the 3D capabilities of AR in science learning, especially in geometry and chemistry [4, 13].

Iordache et al. [11] analyzed the extent to which the specific capabilities of AR are supporting Chemistry learning. The results showed and found out that the interaction paradigm “learning with guidance” has a positive and significant influence on the learning effectiveness and efficiency.

In a recent work [14], a multidimensional model of the perceived quality of AR-based learning has been proposed that has three dimensions: ergonomic, learning, and hedonic quality. The learning quality has two facets: perceived efficiency and perceived usefulness. While the former is measuring the faster and better understanding, the latter refers to the general outcomes of learning in an AR environment.

Gender differences

Yang & Chen [18] analyzed the gender differences in different types of spatial ability and noticed that the differences are mainly related to the spatial perception, mental rotation and spatial visualization of three-dimensional objects.

Their study found gender differences in the performance of a digital game but the pretest and posttest results showed that the differences are smaller after the test which means that the computer-based games may improve the spatial abilities of students.

In a similar vein, the study of Kaufmann [13] noticed that although there are gender differences in AR-based science learning, these are reduced after training, which suggests that computer-based learning software is a useful aid in improving the students' spatial abilities.

More recently, Cheng [5] found out that there are no gender differences in the conceptions of learning science by AR. However, he argued that gender may mediate the relationship between students' epistemic beliefs about scientific knowledge and their conceptions, in the context of AR-based learning.

EMPIRICAL STUDY

Model and variables

The perceived Chemistry learning quality (Q-EDU) is a focal construct which is measured with a set of causal indicators (ARF) pointing to specific AR features. Meantime, the learning quality manifests on two facets: perceived efficiency (PEF) and perceived usefulness (PU). The model is presented in Figure 1.

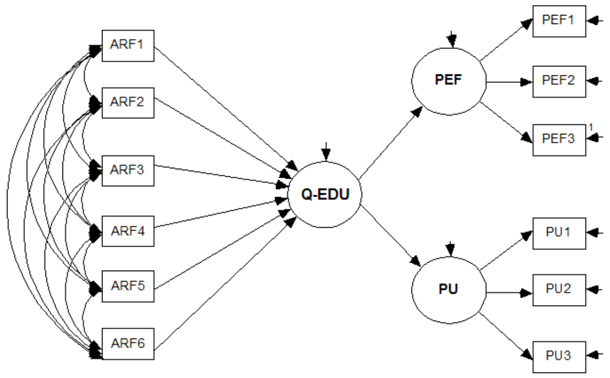

Figure 1. Research model

Overall, it is a formative model with 6 causal indicators and two reflectively measured constructs (PEF and PU). The constructs and measures are presented in Table 1.

Table 1. Constructs and measures

No.	Variables
	AR capabilities (ARF)
ARF1	The augmentation helps to understand the chemical structure of an atom
ARF2	Building a molecule from atoms helps to understand Chemistry
ARF3	Simulating a Chemical reaction with ARTP helps to understand it better
ARF4	Using ARTP helps to understand the periodic table
ARF5	Vocal explanations help interacting with ARTP
ARF6	Interacting with colored balls symbolizing atoms is a good idea
	Perceived efficiency (PEF)
PEF1	ARTP would help me to understand the lesson faster
PEF2	ARTP would help me to learn more quickly
PEF3	ARTP would help me to understand the lesson better
	Perceived usefulness (PU)
PU1	After using ARTP my Chemistry knowledge will improve
PU2	ARTP exercises are useful to test my knowledge
PU3	ARTP helps learning Chemistry

The research model proposed in this study is theoretically grounded. There are many studies arguing on the educational effects of specific AR features [3, 4, 6, 17]. More specifically, the effects of the ARTP capabilities on Chemistry learning have been presented and analyzed on smaller samples, in previous works [11, 12].

The model is testing the hypothesis that the specific features of the learning scenario have a positive and significant influence on the perceived efficiency (ARF $\rightarrow$ PEF) and perceived usefulness (ARF $\rightarrow$ PU).

Equipment, participants and tasks

The Chemistry learning scenario has been implemented on the ARTP that ARTP is a "seated" AR environment: users are looking to a see-through screen where virtual images are superimposed over the perceived image of a real object placed on the table [16].

Figure 2. Building a H2O molecule on ARTP

The implemented learning paradigm was "building with guidance". Two real objects are used in this scenario: a periodic table of chemical elements and four sets of colored balls symbolizing atoms. A remote controller Wii Nintendo has only been used as interaction tool for confirming a selection.

The test was conducted on the ICI's platform which is equipped with 4 ARTP modules. A total number of 186 students (13-15 years old), from which 96 boys and 90 girls tested the platform. After testing, the students were asked to answer a questionnaire by rating the items on a 5-point Likert scale.

The students were assigned three tasks: create atoms by placing a colored ball on the symbol of a chemical element in the periodic table, create molecules and simulate a chemical reaction by using the molecules created in the previous task.

Results

Model validation

Data analysis for checking outliers and normality has been carried out using the SPSS 16.0 for Windows. The analysis showed moderate deviations from normality for all variables. Then a Confirmatory Factor Analysis (CFA) using AMOS for Windows [1] has been carried out to empirically validate the model.

The formative indicators have been validated following the recommendations of Diamantopoulos et al. [8] and Bollen [2]. The collinearity has been analyzed with the

VIF statistic (Variation Inflation Factor) was below the recommended value of 3 [8].

The reflectively measured constructs have been validated for unidimensionality, internal consistency of the scale, and convergent validity, following the recommendations from the literature [10].

Invariance analysis

Since the variables under consideration are measures of an underlying model, an invariance analysis is needed in order to analyze if the respondents are interpreted the variables in the same way. Otherwise, the conclusion could be ambiguous if not erroneous [15].

An invariance analysis has been carried on according to the recommendations from [7, 15]. Since the perceived Chemistry learning quality is formatively measured, the recommendations from [9] have also been considered. However, since the purpose is to assess the gender differences for the causal indicators, apart from checking the slope invariance (structural weights) and the error factor (structural residual) invariance, the structural means invariance has been also examined.

The first step was to split the sample and to test the model for each gender. The model testing results are presented in Table 2 and provide evidence for the configural invariance (same pattern of free and fixed factor loadings on the items).

Table 2. Model testing results for the two subsamples

	N	χ^2	DF	χ^2/DF	TLI	CFI	RMSEA
M	96	59.41	37	1.606	.905	.946	.080
F	90	57.57	37	1.556	.935	.964	.079

The standardized regression coefficients (γ) and the statistical significance (p) of the causal indicators for the two samples are presented in Table 3.

Table 3. Validation results for the causal indicators

Gender		ARF1	ARF2	ARF3	ARF4	ARF5	ARF6
M	γ	0.27	0.21	0.25	0.18	0.17	0.21
	p	0.003	0.029	0.006	0.034	0.070	0.015
F	γ	0.12	0.19	0.29	0.31	0.18	0.31
	p	0.103	0.023	0.000	0.000	0.004	0.000

Except for two cases of marginal significance (ARF5 in the boys' sample and ARF1 in the girls' sample) all indicators are statistically significant.

Then, a multi-group CFA (MGCFA) using AMOS for Windows [1] has been conducted twice, firstly to check the invariance of reflective measures and secondly to include the invariance of formative measures. MGCFA is based on testing a hierarchical series of nested models, starting with a baseline model that fits all the samples together. The parameters are freely estimated and a baseline chi-square value is derived.

The unconstraint model testing (only the right part of the model in Figure 2) showed a good fit of the model with the data: χ^2 = 43.82, p=.038, DF=29, χ^2/DF=1.511, TLI=.970, CFI=.971, RMSEA= .053, srmr =0.0431.

The metric invariance of Q-EDU has been tested by constraining the loadings to be equivalent. The results show a nonsignificant chi-square difference ($\Delta\chi^2$ = 3.948, ΔDF = 4 p = .413), therefore the model exhibits metric invariance. This means that the model has been perceived in the same way in each group.

Testing the scalar invariance is done by constraining the intercepts to be equivalent. The model comparison shows a nonsignificant chi-square difference ($\Delta\chi^2$ = 2.606, ΔDF = 6, p = 0.856, so the model of the perceive learning quality exhibits scalar invariance.

Then the MGCFA has been carried out again, for the entire model. The unconstraint model testing showed a good fit of the model with the data: χ^2 = 116.981, p=.001, DF=74, χ^2/DF=1.581, TLI=.922, CFI=.956, RMSEA= .056, srmr =0.0493.

The test for the structural weights, which includes the test for the slope invariance, resulted in a non-significant chi-square difference ($\Delta\chi^2$ = 6.36, ΔDF = 7, p = 0.498), which means that the causal indicators are invariant across gender. The results of the invariance analysis are presented in Table 4.

Table 4. Invariance analysis results

Model	DF	χ^2	CFI	Δ DF	Δ χ^2	Δ CFI	P
Unconstraint	74	116.98	0.956				
Meas. weights	78	119.65	0.958	4	2.67	0.002	0.614
Meas. intercepts	84	125.08	0.958	6	5.42	0.000	0.490
Struct. weights	91	131.45	0.959	7	6.36	0.001	0.498
Struct. means	97	138.03	0.958	6	6.58	-0.001	0.361
Struct covariance	118	163.97	0.953	21	25.94	-0.005	0.209
Struct residuals	121	170.39	0.95	3	6.41	-0.003	0.093

Since the model has metric, scalar, slope, structural means, and factor error invariance, it is possible to compare the contribution of causal indicators, the observed scores of the reflective and formative measures as well as the means of the latent variables PEF and PU.

Gender differences

The gender differences in the perceived learning quality are presented in Table 5.

Table 5. Mean values for PEF and PU

Gender	PEF1	PEF2	PEF3	PU1	PU2	PU3
M	4.13	4.13	4.02	3.71	3.97	4.17
F	4.06	3.82	4.02	3.62	3.94	4.20
Total	4.10	4.02	4.02	3.68	3.96	4.18

The male students scored higher all items except for PU3 (ARTP helps learning Chemistry) and PEF3 (ARTP helps to understand the lesson better). However, a one-way ANOVA test shows that the differences are not statistically significant.

At construct level, the results show that male students have a higher perception of both the perceived efficiency (M=4.09 vs. F=3.96) and perceived usefulness (M=3.95 vs. F=3.92). The gender differences are not statistically significant.

The gender differences for the mean values of the causal indicators are presented in Table 6.

Table 6. Mean values for the causal indicators

Gender	ARF1	ARF2	ARF3	ARF4	ARF5	ARF6
M	4.04	4.20	4.23	4.20	3.99	4.39
F	4.02	4.07	4.10	4.40	4.10	4.40
Total	4.03	4.13	4.17	4.30	4.04	4.39

The boys have a higher perception of the first three indicators: augmentation, building with guidance, and simulating chemical reactions. The girls scored higher the use of real objects and the vocal explanations.

There are gender differences as regards the contribution of each indicator to the perceived learning quality (see Table 3). For the boys, the most useful AR features are the augmentation and the simulation of a chemical reaction. For the girls, the most useful features are the real objects (periodic table and colored balls), and the simulation of a chemical reaction.

CONCLUSION

This work is contributing with a broader perspective on the perceived learning quality of an AR-based platform for learning Chemistry. The inclusion of causal indicators makes it possible to assess how it manifests on the perceived learning efficiency and perceived usefulness as well as how the specific AR features contribute to the learning quality.

Overall, the students considered that the AR technology helps understanding the atom structure and the chemical reactions. The outcomes of the perceived learning quality are the efficiency and usefulness of the learning process.

The results showed that the model is invariant across genders thus enabling comparison. Although the gender differences are not statistically significant as regards the observed scores, there are several interesting differences as regards the contribution of the AR features. While for the boys, the augmentation (computer generated image) was the most useful feature, for the girls using the real objects (colored balls and periodic table) were the most useful. For both genders, simulating a chemical reaction had an important contribution to the perceived learning quality.

Acknowledgement

The AR platform and the Chemistry application were developed in the European research project ARiSE (FP6 027039).

REFERENCES

1. Arbuckle J.L. (2006). *AMOS 7.0 User's Guide*. Amos Development Corporation
2. Boolen K. (2011) Evaluating effect, composite and causal indicators in structural equation models. *MIS Quarterly 35(2)*, 359-372.
3. Cai S., Wang X., Chiang FK. (2014) A case study of Augmented Reality simulation system application in chemistry course. *Computers in Human Behavior, 37*, 31-40.
4. Chen Yu-Chien (2006). A study of comparing the use of augmented reality and physical models in chemistry education. *ACM International Conference on Virtual Reality Continuum and Its Applications*, 369-372.
5. Cheng K.H. (2018) Surveying students' conceptions of learning science by Augmented Reality and their scientific epistemic beliefs. *Eurasia Journal of Science, Technology, and Education, 14(4)*, 1147-1159.
6. Cheng K.H., Tsai C.C. (2013). Affordances of augmented reality in science learning: suggestions for future research. *Journal of Science Education and Technology*, *22*(4), pp. 449-462
7. Cheung GW and Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modelling 9(2)*, 233-255.
8. Diamantopoulos A., Riefler P., Roth K. (2008) Advancing formative measurement models. *Journal of Business Research 61*, 1203-1218.
9. Diamantopoulos A, Papadopoulos N (2010) Assessing the cross-national invariance of formative measures: guidelines for international business researchers. *Journal of International Business Studies*, *41*, 360-370.
10. Hair J.F., Black W.C., Babin B.J., Anderson R.E., Tatham, R.L. (2006). Multivariate Data Analysis. 6th Ed., Prentice Hall, 2006.
11. Iordache D.D., Pribeanu C., Balog A. (2012). Influence of Specific AR Capabilities on the Learning Effectiveness and Efficiency. *Studies in Informatics and Control*, *21*(3), 233-240
12. Iordache DD, Pribeanu C (2013) How does the Augmented Reality help the understanding of Chemistry: A model based on causal indicators. Stefanut D, Rusu C. (Eds.) *Proc. RoCHI 2013,* MatrixRom, 29-34
13. Kaufmann H (2004) *Geometry education with Augmented Reality*. PhD Thesis, Wienna University of Technology, 179p.
14. Pribeanu C, Balog A, Iordache DD (2017) Measuring the perceived quality of an AR-based learning application: a multi-dimensional model. *Interactive Learning Environments*, 25(4), 482-485.
15. Vandenberg R.J., Lance C.E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods 3(1*), 4-70.
16. Wind J., Riege K., Bogen M. (2007). Spinnstube®: A Seated Augmented Reality Display System. *Virtual Environments: Proc. IPT-EGVE – EG/ACM Symposium*, 17-23.
17. Wu H. K., Wen-Yu Lee S., Chang H. Y., Liang J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education 62*, 41-49.
18. Yang JC, Chen S (2010) Effects of gender differences and spatial abilities within a digital pentominoes game. *Computers & Education*, *55*, 1220-1233.