



Teaching empathy to nursing students: A randomised controlled trial

Pilar Bas-Sarmiento^{a,*}, Martina Fernández-Gutiérrez^a, Mercedes Díaz-Rodríguez^b, iCARE Team^c (Concepción Carnicer-Fuentes, Cristina Castro-Yuste, María José García-Cabanillas, Cristina Gavira-Fernández, María de los Ángeles Martelo-Baro, Olga Paloma-Castro, María del Carmen Paublete-Herrera, María Jesús Rodríguez-Cornejo, Luis Moreno-Corral)

^a Faculty of Nursing, University of Cádiz, Venus Street, 11207 Algeciras, Cádiz, Spain

^b Faculty of Nursing and Physiotherapy, University of Cadiz, Ana de Viya Avenue 52, 11009 Cádiz, Spain

^c Nursing and Physiotherapy Department, University of Cádiz, Ana de Viya Avenue 52, 11009 Cádiz, Spain

ARTICLE INFO

Keywords:

Empathy
Health communication
Intervention study
Nursing education
Nursing students
Professional-patient relations
Randomised controlled trial
Role-playing

ABSTRACT

Background: Empathy has been reported to produce a positive effect on improving patient health outcomes, becoming a fundamental skill in any health personnel-patient relationship.

Objective: To evaluate the effectiveness of an intervention designed to improve the nursing students' empathy, the learning perception, the improvement of the perception in the understanding of the content, and in the degree of difficulty as well as the acquisition of skills.

Design: Multicentre randomised controlled trial.

Setting: This research was conducted at two schools of nursing at a public university in the Southwest of Spain.
Participants: 116 nursing students were randomly assigned to an experimental or a control group (delayed intervention group once the post-training analysis was completed) during the second semester of the 2015/2016 academic year.

Methods: Pre-test, post-test, and follow-up data were obtained for each group using a simulated clinical interview. Empathy was the primary outcome (The Consultation and Relational Empathy Measure, Jefferson Scale of Empathy student version, Reynolds Empathy Scale, and Carkhuff Scale). The students' perceived knowledge, the learning perception, the self-esteem (Rosenberg Self-Esteem Scale) and the understanding of the content and acquisition of skills (ad-hoc questions) were also analysed.

Results: The results were improved in all the measures conducted in the experimental groups at the different centres after the intervention. The mean post-test simulation scores were higher than the pre-test with statistically significant differences. The results were maintained in the follow-up. The student's perception of learning and the perception of understanding of the content and the acquisition of skills were improved as well.

Conclusion: The study support that training in empathic competence is effective.

1. Introduction

Empathy is both a multidimensional and skills-based construct according to the conceptual background provided by Mercer and Reynolds (2002). They described four components of the empathy construct: cognitive elements (identify and understand another's feelings), emotional (experiment with and share feelings), moral (internal encouragement to empathise), and relationship (communicative response to understanding).

The investigation linked empathetic care with better patient training and improvement of the health results (Batt-Rawden et al.,

2013; Bikker et al., 2015; Kelm et al., 2014). It has also been noted that empathy improves the quality of information that is communicated between the professional and patients, as well as the quality of care, the patient and personnel satisfaction (Chen and Forbes, 2014), and the participation and therapeutic adherence of the patient (Batt-Rawden et al., 2013; Kelm et al., 2014). It also lowers the chances of misunderstandings, and, as such, demands (Kelm et al., 2014; Nosek et al., 2014; Williams et al., 2015). Empathy even minimises the use of sanitary resources and thus, expenditures (Kelm et al., 2014).

Despite the aforementioned, recent studies show a decline in empathy as much in students and professionals as an increase in nursing

* Corresponding author at: Faculty of Nursing, University of Cadiz, Venus Street, s/n, 11207, Algeciras, Cádiz, Spain.

E-mail addresses: pilar.bas@uca.es (P. Bas-Sarmiento), martina.fernandez@uca.es (M. Fernández-Gutiérrez), mercedes.diaz@uca.es (M. Díaz-Rodríguez), luis.moreno@uca.es (L. Moreno-Corral).

<https://doi.org/10.1016/j.nedt.2019.06.002>

Received 5 September 2018; Received in revised form 13 May 2019; Accepted 10 June 2019

0260-6917/ © 2019 Elsevier Ltd. All rights reserved.

(Díaz-Narváez et al., 2017; Nosek et al., 2014) and medical practices (Chen and Forbes, 2014; Kelm et al., 2014; Neumann et al., 2012; Ward, 2016). It has also been noted that medical students do not feel comfortable empathising with their clinical trial patients (Lobchuk et al., 2016).

That is why instructors of empathy are vital to health students such that, in the future, health professionals will consider empathy as an essential skill (Batt-Rawden et al., 2013; Kelm et al., 2014; Ward et al., 2012; Williams and Stickley, 2010). As such, it is absolutely necessary to develop an adequate level of empathy by effective instruction based on scientific evidence.

In a prior study (Bas-Sarmiento et al., 2017), we carried out an empathy intervention with a quasi-experimental design, which achieved promising results. In order to compensate the methodological limitations of this type of design, we have proposed to evaluate, in a randomised multicentre controlled trial, the benefits of empathy training available to nursing students. The specific objectives were: (1) to evaluate the effects of empathy training, the level of the students' learning perception, the improvements of the perception in the understanding of the content and in the degree of difficulty as well as the acquisition of skills, and the level of self-esteem and (2) to evaluate whether the effects of this training are independent of the learning centres at which they were taught and if they withstand the test of time.

2. Methods

2.1. Study design

This study was a multicentre open randomised controlled trial, masked for data analysis. The participants students were randomly assigned to either an immediate intervention group (experimental group 1: participation in the empathy training) or a waitlisted group (control group/experimental group 2: delayed intervention).

This type of design was named “modified cohort-control design” by McKillip (1979), and it allowed the collection of pre- and post-intervention measurements from two separate groups of students, where the intervention for group 2 lagged behind that for group 1. This design controlled for the effects of exposure to the standard curriculum and the passage of time on the dependent variables. Since the waitlisted students eventually received the intervention, we replicated the data collection and intervention in the two separate groups and the two different centres (Fig. 1).

2.2. Setting and participants

The sample was composed of second-year nursing students at the University of Cadiz who were enrolled in an Interpersonal

Communication Skills class. The training was part of the scheduled teaching activities. During the subject presentation (coinciding with the pre-test evaluation), the students were informed of the training objectives, schedule, and methodology.

The intervention was performed during the second semester of the 2015/2016 academic year at two nursing faculties of the University of Cadiz located in two cities 100 km apart (the Nursing Faculty in Algeciras — Centre 1 and the Nursing and Physiotherapy Faculty of Jerez — Centre 2).

2.3. Randomised

A computer-generated randomisation list was used to assign the students to one of the two groups by a research assistant who didn't participate in the study (Fig. 1). This information was available only for the principal investigator, who assigned the students a sequential number that was placed in the class list. Half of the participants of each centre were randomly assigned to the experimental and control groups, respectively.

2.4. Intervention

The training was accomplished by the university teaching personnel with a specific training in communication (a graduate and a PhD in psychology, a graduate in nursing and a PhD in health sciences, and a physician and a PhD in medicine). A group of students was assigned to each professor with the aim of controlling the influence of strange variables, academic training, and/or the professors' ability. The intervention consisted of two weekly meetings of 2 h each with seven sessions (14 h). After some literature recommendations concerning intermediate duration interventions (Hart et al., 2006) as well as the results obtained from a previous study (Bas-Sarmiento et al., 2017), there was a reduction in the length of the initial project, from 20 to 14 h (from 10 to 7 sessions). The intervention is detailed in Supplementary data 1.

2.5. Outcome measures and instruments

2.5.1. Main outcome: empathy

The general measure of the empathy level of the students was measured by the Spanish adaptation of Díaz et al. (2014) of the Jefferson Scale of Empathy (JSE) (Hojat et al., 2001). This measure allows to ensure the homogeneity of the sample at the baseline.

During the simulations, the information was triangulated by measuring the empathic competency perceived by the student (health personnel), the empathy perceived by the patient during the student's performance, and the observation of the behaviour by independent observers:

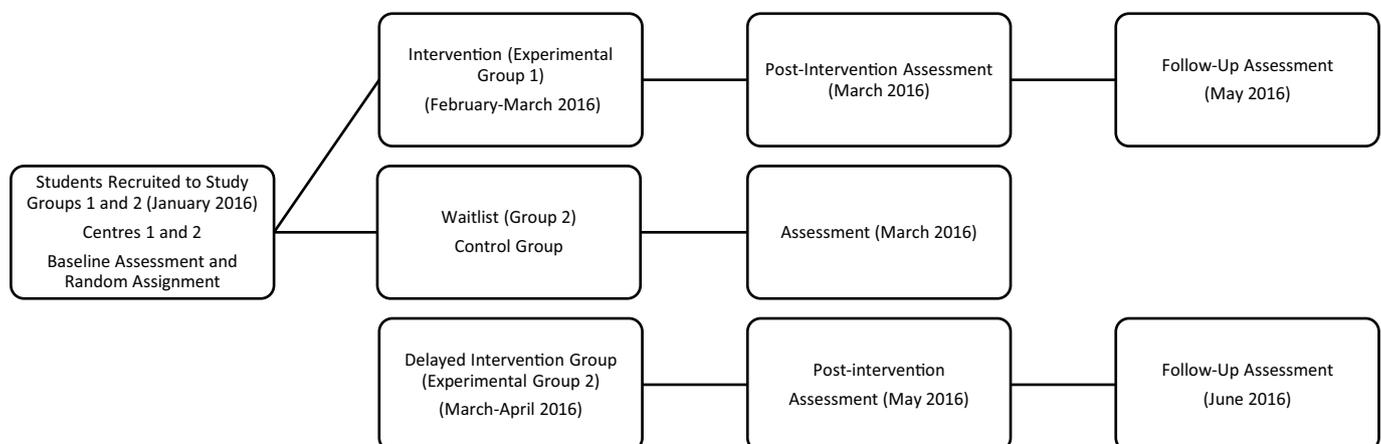


Fig. 1. Study design.

The behaviours and attitudes of the students during an oral interaction with a patient (simulated clinical interviews) were evaluated by the Spanish version of the Reynolds Empathy Scale (Bellver-Navalón et al., 2004; Reynolds, 2000). The scale has good psychometric properties.

The Consultation and Relational Empathy (CARE) Measure (Mercer et al., 2004) was used to measure the professional's empathic competency from the point of view of patients. This measurement has been adapted to the Spanish setting with the same reliability (internal consistency) as the original scale.

The empathic competency of the students was also assessed by independent observers (nurses and/or university professors who did not participate in the training) using a rating scale pre-established in the research protocol following the Carkhuff Scale (Carkhuff, 1969) as the gold standard. Independent observers were blinded for data analysis.

2.5.2. Secondary outcomes

The students' perception of their learning was evaluated after the training using a 10-point Likert scale. The students' assessments of the degree of difficulty in understanding the contents and the acquisition of empathic competence compared to the degree of difficulty that they finally perceived and the extent to which training favoured the understanding of the contents and the acquisition of competencies were measured using an ad-hoc 5-point Likert scale (see in Supplementary data 2).

Finally, the Spanish version of the Rosenberg Self-Esteem Scale which was validated in Spanish University Students was used (Martín-Albo et al., 2007).

2.6. Data collection

The evaluation was performed using simulated clinical interviews of 15-minute duration with amateur actors from the School of Theatre who voluntarily participate and were previously trained to act as patients. Each student was offered three opportunities to be empathic: (1) the patient verbalises that he/she is very nervous because a family member suffered serious consequences from the same pathology; (2) the patient asks, with a worried gesture, if he/she is going to heal; and (3) the patient shows his/her discomfort and/or inability to abandon an unhealthy habit that is a risk factor.

Prior to the simulation, the students received documentation regarding the clinical case and the necessary instructions for the course of the interview in which, among other things, they were expressly asked to empathise.

After the simulated clinical interview, the students were asked to complete the Reynolds Empathy Scale, satisfaction with the learning survey, and the Rosenberg Self-Esteem Scale. The patient (actor) was also asked to assess each student using the Consultation and Relational Empathy (CARE) Measure.

The follow-up evaluation was conducted one month after the intervention. The follow-up of experimental group 1 was scheduled to coincide with the post-intervention measurement of experimental group 2 (the delayed group). The follow-up of the delayed group (experimental group 2) was developed in the context of the Structured Objective Evaluation of Nursing Competencies (EOECE), corresponding to the Practicum II subject, which consists of evaluating the competencies acquired by the student at four clinical simulation stations (cardiovascular, medication administration, skin care, and the nursing process). At one of the four stations (cardiovascular consultation), the simulation was incorporated for the evaluation of empathy following the same procedure described. The EOECE was not carried out at Centre 2, so the follow-up of the delayed group at the Centre 2 was not possible.

2.7. Analysis data

Prior to the analysis, the normality of the variables was evaluated through the Kolmogorov-Smirnov and Levene test. Continuous variables were expressed as summary measures (means, medians) and measure of dispersion (standard deviation, range) and the categorical variables were expressed frequency and percentage. The differences in baseline of characterization variable of the sample between intervention and control groups were compared by the non-parametric statistics U Mann-Whitney test. The differences between qualitative variables were compared using Fisher test. The differences between the efficacy variables, in this case continuous variables, were compared by U Mann-Whitney test. For the intergroup analysis, the Student-Fisher *t*-test for paired data or the Wilcoxon signed-rank test was used in order to observe differences in the case of repeated measurements in response variables.

Effect sizes were used (Cohen's conventional criteria) to complement statistical hypothesis testing, and play an important role in power analyses. Spearman's correlation was used to analyse the strength of the linear relationship between the paired data. All *p*-values were based on a two-tailed distribution and a statistical significance of 0.05. The statistical analysis was carried out with the assistance of the SPSS program version 22.0 for Windows (IBM).

2.8. Ethical issues

This work was conducted in accordance with the Declaration of Helsinki. The right to anonymity and the confidentiality of the participants were protected at all times. The sessions were recorded with the informed consent of the participants.

3. Results

A flow chart of the students' progress through the trial is shown in Fig. 2. Finally, 116 out of 122 enrolled students were included in the study. The average age was 22.78 (SD = 5.43) and 86 were women (74.1%). Overall, 72 students belonged to the Nursing Faculty (Centre 1) and 44 students belonged to the Nursing and Physiotherapy Faculty (Centre 2) at the University of Cadiz. The average age at Centre 1 was 23.38 (SD = 6.36), whereas in Centre 2 it was 21.81 (SD = 3.24). Focusing on gender, 69.4% of Centre 1's students were women (*N* = 50) vs 81.8% at Centre 2 (*N* = 36).

No significant differences in demographics were found comparing the intervention and control groups. The baseline scores were similar between the intervention and control groups (see in Supplementary data 3). Thus, the empathy general measurement (Jefferson Scale of Empathy) as well as the rest of the parameters indicated homogeneity between the groups at baseline.

Tables 1, 2, and 3 show descriptive statistical (mean and standard deviation) and statistical significance between the scores pre- and post-training, as well as the results of monitoring both centres in different groups (experimental groups 1 and 2 and the control group). At the Centre 1, there were improvements in every evaluated measurement, and the pre-training/post-training differences were statistically significant, in GE1 as well as in GE2 with the exception of the Jefferson Empathy Scale scores in GE2; although they showed an improvement, they were not statistically significant. In Control Group (GC), there were no statistically significant differences in the evaluated parameters as expected, which established the training's effectiveness. Comparing the students' post-training scores to those in the monitoring, the results were maintained. In the same way, if the students' baseline scores (pre-training) were compared to the scores during the monitoring (GE1 and GE2), there was evident improvement in the results; statically significant differences appeared in the three evaluated parameters during the simulation. In center 1, there were statistically significant differences between the scores of control and experimental-1 groups for all

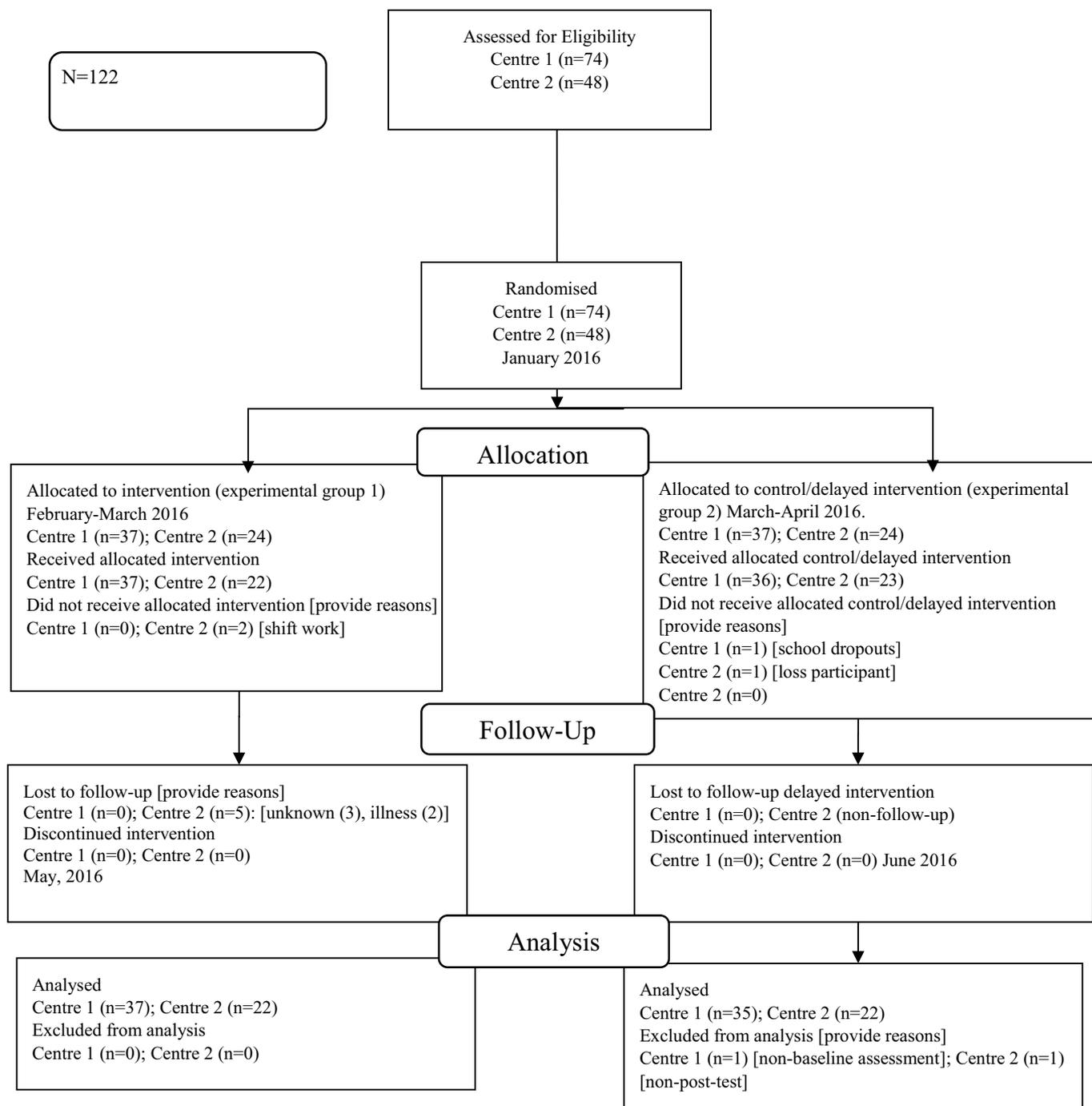


Fig. 2. Flow chart of the participants and the recruitment and follow-up periods.

the main variables, except Jefferson. Between control and experimental-2 groups results were similar, except Reynolds score improved but was not statistically significant.

In the same way, at the Centre 2, there were improvements in every evaluated measure during the simulation, with statically significant differences in both pre- and post-training in GE1 and GE2 (with the exception of the Jefferson Empathy Scale scores; although they showed an improvement, they were not statistically significant). The GE1 maintained the level of perceived learning during the monitoring as well as improved in some scores regarding the post-test. Nevertheless, the control group in Centre 2 showed special behaviour, improving post-training scores with statically significant results for the observers and the students. In fact, in center 2, there were statistically significant

differences between the scores of the control and experimental groups for the evaluation of Carkhuff and CARE (only between the control group and the experimental group 2).

Although the male representation was considerably inferior, in general, there were no statically significant differences found between the men and women at any of the centres. At Centre 1, there were significant differences in self-esteem levels between the men and women, and the average rank was inferior in the women in the post-tests (30.28 vs 33.50; $z = -2.599, p = 0.009$) as well as in monitoring (31.93 vs 35.22 for the men; $z = -2.750, p = 0.006$). Age correlated negatively with the general empathy level (Jefferson Empathy Scale score) in GE1 Centre 1 at baseline ($r = 0.363, p = 0.015$); this was not the case in the rest of groups and measurements.

Table 1
Mean, standard deviation, differences between groups, and Wilcoxon signed-rank test for pre- and post-educational intervention according to the group and education centre.

	CG-C1				EG1-C1				EG2-C1				CG-C2				EG1-C2				EG2-C2				BETWEEN GROUP ANALYSIS				
	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD	Z(p)	Z(p)	
Carkhuff	Pre	35	1.87 ± 0.77	37	1.81 ± 0.67	35	1.87 ± 0.77	22	1.56 ± 0.72	21	1.57 ± 0.69	22	1.56 ± 0.72	22	1.57 ± 0.69	22	1.56 ± 0.72	22	1.56 ± 0.72	22	1.56 ± 0.72	22	1.56 ± 0.72	22	1.56 ± 0.72	CG-EG1: -2.548 (0.011)	CG-EG2: -4.548 (0.000)		
	Post	35	2.00 ± 0.50	37	4.24 ± 0.74	35	4.24 ± 0.58	22	2.15 ± 0.96	22	2.88 ± 0.88	22	2.15 ± 0.96	22	2.88 ± 0.88	22	2.88 ± 0.88	22	2.15 ± 0.96	22	2.15 ± 0.96	22	2.15 ± 0.96	22	2.15 ± 0.96	CG-EG1: -7.278 (0.000)	CG-EG2: -7.267 (0.000)		
RES	Pre	35	75.31 ± 11.47	37	77.51 ± 14.69	35	75.31 ± 11.47	22	76.06 ± 12.95	22	73.97 ± 13.84	22	76.06 ± 12.95	22	73.97 ± 13.84	22	76.06 ± 12.95	22	76.06 ± 12.95	22	76.06 ± 12.95	22	76.06 ± 12.95	22	76.06 ± 12.95	CG-EG1: -2.631 (0.009)	CG-EG2: -1.721 (0.085)		
	Post	35	74.70 ± 15.57	37	84.70 ± 15.62	35	81.54 ± 10.80	22	84.43 ± 12.47	22	88.06 ± 12.63	22	84.43 ± 12.47	22	88.06 ± 12.63	22	88.06 ± 12.63	22	84.43 ± 12.47	22	84.43 ± 12.47	22	84.43 ± 12.47	22	84.43 ± 12.47	CG-EG1: -3.440 (0.001)	CG-EG2: -2.500 (0.012)		
CARE	Pre	35	23.17 ± 8.81	37	24.89 ± 7.38	35	23.17 ± 8.81	22	22.50 (± 7.86)	22	19.81 ± 5.40	22	22.50 ± 7.86	22	19.81 ± 5.40	22	22.50 ± 7.86	22	22.50 ± 7.86	22	22.50 ± 7.86	22	22.50 ± 7.86	22	22.50 ± 7.86	CG-EG1: -4.024 (0.000)	CG-EG2: -4.257 (0.000)		
	Post	35	23.11 ± 5.88	37	29.00 ± 5.98	35	28.82 ± 5.16	22	27.63 (± 8.12)	22	31.22 ± 11.62	22	27.63 ± 8.12	22	31.22 ± 11.62	22	31.22 ± 11.62	22	27.63 ± 8.12	22	27.63 ± 8.12	22	27.63 ± 8.12	22	27.63 ± 8.12	CG-EG1: -3.541 (0.000)	CG-EG2: -2.980 (0.003)		
Jefferson	Pre	35	115.11 ± 8.96	37	116.72 ± 8.58	35	115.11 ± 8.96	22	108.09 ± 13.01	22	110.18 ± 11.84	22	108.09 ± 13.01	22	110.18 ± 11.84	22	108.09 ± 13.01	22	108.09 ± 13.01	22	108.09 ± 13.01	22	108.09 ± 13.01	22	108.09 ± 13.01	CG-EG1: -1.083 (0.279)	CG-EG2: -0.006 (0.995)		
	Post	35	116.77 ± 10.83	37	119.94 ± 7.77	35	116.51 ± 9.51	22	113.50 ± 11.48	22	110.36 ± 18.10	22	113.50 ± 11.48	22	110.36 ± 18.10	22	111.13 ± 8.97	22	113.50 ± 11.48	22	113.50 ± 11.48	22	113.50 ± 11.48	22	113.50 ± 11.48	CG-EG1: -0.247 (0.085)	CG-EG2: -1.022 (0.307)		
Effect size (d)	Pre		-0.655 (0.512)		-3.028 (0.002)		-2.914 (0.004)		-1.853 (0.064)		-3.541 (0.000)		-1.853 (0.064)		-3.541 (0.000)		-4.043 (0.000)		-1.853 (0.064)		-1.853 (0.064)		-1.853 (0.064)		-1.853 (0.064)				
	Post		-0.672 (0.502)		-2.284 (0.022)		-1.144 (0.252)		-1.447 (0.148)		-0.522 (0.602)		-1.447 (0.148)		-0.522 (0.602)		-0.262 (0.794)		-1.447 (0.148)		-1.447 (0.148)		-1.447 (0.148)		-1.447 (0.148)				

CG-C1 = Control Group-Center1; EG1-C1 = Experimental Group 1-Center 1; EG2-C1 = Experimental Group 2-Center 1; CG-C2 = Control Group-Center2; EG1-C2 = Experimental Group 1-Center 2; EG2-C2 = Experimental Group 2-Center 2.

Table 2
Mean, standard deviation, and Wilcoxon signed-rank test for post-educational intervention/follow-up, according to the group and education centre.

Group	Variables	Centre	N	Mean ± SD	z (p)	Effect size (d)	Centre	N	Mean ± SD	z (p)	Effect size (d)	
Experimental 1	Carkhuff (Post-I)	1	37	4.24 ± 0.74	-0.166 (0.868)		2	22	2.88 ± 0.88	-2.789 (0.005)	0.59	
	Carkhuff (Follow-up)	1	37	4.25 ± 0.50			2	17	3.79 ± 0.63			
	RES (Post-I)	1	37	84.70 ± 15.62	-0.045 (0.964)		2	22	88.06 ± 12.63	-0.829 (0.407)		
	RES (Follow-up)	1	37	84.85 ± 12.08			2	17	89.73 ± 11.96			
	CARE (Post-I)	1	37	29 ± 5.98	-0.505 (0.614)		2	22	31.22 ± 11.62	-2.393 (0.017)	0.51	
	CARE (Follow-up)	1	37	29.70 ± 6.52			2	17	40.94 ± 7.58			
	Jefferson (Post-I)	1	37	119.94 ± 7.77	-1.890 (0.059)		2	22	110.36 ± 18.10	-0.735 (0.463)		
	Jefferson (Follow-up)	1	37	117.91 ± 8.28			2	17	112.56 ± 11.61			
	Self-esteem (Post-I)	1	37	31.05 ± 5.39	-2.561 (0.010)	0.42	2	18	34.94 ± 4.37	-0.701 (0.483)		
	Self-esteem (Follow-up)	1	37	32.83 ± 4.74			2	17	34.94 ± 5.05			
	Learning (Post-I)	1	34	7.44 ± 1.90	-0.973 (0.330)		2	17	7.05 ± 1.85	-2.360 (0.018)	0.57	
	Learning (Follow-up)	1	35	7.48 ± 1.66			2	16	8.18 ± 0.98			
	Experimental 2 (Follow-up: EOECE)	Carkhuff (Post-I)	1	35	4.24 ± 0.58	-1.629 (0.103)		2				
		Carkhuff (Follow-up)	1	35	3.94 ± 0.69			2				
		RES (Post-I)	1	35	81.54 ± 10.80	-0.427 (0.669)		2				
RES (Follow-up)		1	35	82.21 ± 12.39			2					
CARE (Post-I)		1	35	28.82 ± 5.16	-3.870 (0.000)	0.65	2					
CARE (Follow-up)		1	35	35.48 ± 7.77			2					
Jefferson (Post-I)		1	35	116.51 ± 9.51	-0.287 (0.774)		2					
Jefferson (Follow-up)		1	35	116.57 ± 8.65			2					
Self-esteem (Post-I)		1	35	31.48 ± 4.44	-3.090 (0.002)	0.52	2					
Self-esteem (Follow-up)		1	34	33.08 ± 4.73			2					
Learning (Post-I)		1	33	7.24 ± 1.41	-1.696 (0.090)		2					
Learning (Follow-up)		1	33	7.63 ± 1.24			2					

Correlations (see Tables 4 and 5) were found between the different post-training scores. After the intervention in the experimental groups at both centres, there was a correlation between the RES scores and the learning scale perceived (Centre 1 GE1: $r = 0.835, p = 0.000$, and GE2: $r = 0.0389, p = 0.013$; Centre 2 GE1: $r = 0.429, p = 0.043$, and GE2: $r = 0.451, p = 0.017$).

In the same way, the learning rank perceived by the GE1 students correlated significantly with the scores in the Carkhuff Scale by the observers at Centre 1 ($r = 0.384, p = 0.012$) and Centre 2 ($r = 0.485, p = 0.024$), with the self-esteem level for the students in both centres (Centre 1: $r = 0.296, p = 0.045$; Centre 2: $r = 0.419, p = 0.047$) and the score provided by the patients (CARE), only in Centre 1 for GE1 ($r = 0.396, p = 0.010$) and GE2 ($r = 0.357, p = 0.021$).

Jefferson's scores, with slight variations, tended to remain stable, obtaining correlations pre-and post-test for both centres in GE1 (Centre 1: $r = 0.544, p = 0.000$; Centre 2: $r = 0.409, p = 0.029$) and in GE2 only at Centre 2 ($r = 0.475, p = 0.013$). After the intervention, the general empathy scoring (Jefferson Empathy Scale) correlated with the patients (CARE) and observers' valuations (the Carkhuff Scale) in GE1 for both centres (CARE Centre 1: $r = 0.447, p = 0.003$; Centre 2:

$r = 0.416, p = 0.027$; Carkhuff Centre 1: $r = 0.393, p = 0.008$; Centre 2: $r = 0.547, p = 0.004$) and with the self-esteem for both groups at Centre 2 (GE1: $r = 0.485, p = 0.021$; GE2: $r = 0.392, p = 0.035$). In the same way, the scores of the Jefferson Empathy Scale and RES correlated at both centres in just one experimental group (Centre 1 GE2: $r = 0.517, p = 0.001$; Centre 2 GE1: $r = 0.545, p = 0.004$).

Significant correlations were noted between the observers and patients in every case but GE1 at Centre 1 (Centre 1 GE2: $r = 0.409, p = 0.007$; Centre 2 GE1: $r = 0.743, p = 0.000$, and GE2: $r = 0.568, p = 0.03$).

Regarding the difficulty rank noted in the contents' comprehension and competence acquisition (see in Supplementary data 4), 43.9% of the students at the beginning of the training found the comprehension and/or the acquisition of the competence too difficult. At the end of the intervention, the students considered the same difficulty in 21.55% of the cases. Overall, 71.55% evaluated that the innovative elements used were favourable for the contents' comprehension and/or the acquisition of the competencies associated with the intervention.

Table 3
Mean, deviation standard, and Wilcoxon signed-rank test pre-educational intervention/follow-up according to the group and education centre.

Group	Variables	Centre	N	Mean ± SD	z (p)	Effect size (d)	Centre	N	Mean ± SD	z (p)	Effect size (d)
Experimental 1	Carkhuff (Pre-I)	1	37	1.81 ± 0.67	-5.332 (0.000)	0.87	2	21	1.57 ± 0.69	-3.639 (0.000)	0.79
	Carkhuff (Follow-up)	1	37	4.25 ± 0.50			2	17	3.79 ± 0.63		
	RES (Pre-I)	1	37	77.51 ± 14.69	-2.587 (0.010)	0.42	2	22	73.97 ± 13.84	-3.575 (0.000)	0.76
	RES (Follow-up)	1	37	84.85 ± 12.08			2	17	89.73 ± 11.96		
	CARE (Pre-I)	1	37	24.89 ± 7.38	-3.097 (0.002)	0.50	2	22	19.81 ± 5.40	-3.623 (0.000)	0.77
	CARE (Follow-up)	1	37	29.70 ± 6.52			2	17	40.94 ± 7.58		
	Jefferson (Pre-I)	1	37	116.72 ± 8.58	-0.869 (0.385)		2	22	110.18 ± 11.84	-0.994 (0.320)	
	Jefferson (Follow-up)	1	37	117.9 ± 8.28			2	17	112.56 ± 11.61		
	Experimental 2	Carkhuff (Pre-I)	1	35	1.87 ± 0.77	-5.026 (0.000)	0.84	2	22	1.56 ± 0.72	-
Carkhuff (Follow-up)		1	35	3.94 ± 0.69	2			-	-		
RES (Pre-I)		1	35	75.31 ± 11.47	-2.503 (0.012)	0.42	2	22	76.06 ± 12.95	-	
RES (Follow-up)		1	35	82.21 ± 12.39			2	-	-		
CARE (Pre-I)		1	35	23.17 ± 8.81	-4.523 (0.000)	0.76	2	22	22.50 ± 7.86	-	
CARE (Follow-up)		1	35	35.48 ± 7.77			2	-	-		
Jefferson (Pre-I)		1	35	115.11 ± 8.96	-0.908 (0.364)		2	22	108.09 ± 13.01	-	
Jefferson (Follow-up)		1	35	116.57 ± 8.65			2	-	-		

4. Discussion

These results support evidence of effectiveness of the experiential training for the acquisition of empathic competency, as noted by other similar studies (Bas-Sarmiento et al., 2017; Grau et al., 2016; Lor et al., 2015; Nosek et al., 2014; Yang et al., 2014). From the perspective of the independent observers, as well as from the patients' view and the students' perception, the scores improved significantly with respect to the baseline. However, it is important to be cautious due to the results of the control group in Centre 2. An explanation might be the possible contamination of the group since in this case, the contents and activities were visible in the Virtual Campus of the course for all of the students and most did the activities that they were not supposed to do as a control group. A repetition of the intervention in other contexts would be interesting in order to confirm its efficacy.

There is a consensus in the literature that empathy levels vary according to the characteristics of the participants, such as gender, age, and self-esteem (Batt-Rawden et al., 2013). Thus, empathy is greater in younger participants (Batt-Rawden et al., 2013) and in women than in men (Ilhan et al., 2016; Nosek et al., 2014). Although in this case there were no statistically significant differences with respect to the scores of the men and women, the reduced male sample size could explain these results. Regarding age, the general empathy scores (Jefferson Empathy Scale) correlated with age in only one group, which may have been influenced by the fact that the majority of the sample was part of the same age group.

As in other investigations (Nosek et al., 2014; Yang et al., 2014), the assessment of the perception of learning and the utility of what was learned was positive. Those who self-evaluated best in empathy during the simulation were those who felt they had learned more. In general, these students obtained the highest self-esteem scores. In this sense, the literature shows how self-esteem is enhanced by training in social skills (Pasli-Gurdogan et al., 2016). In the same way, it seems that the learning perceived by the students was real given that it coincided with the highest score on the part of the observers and of the patients themselves who were, in effect, those who received the care.

In agreement with a prior study (Bas-Sarmiento et al., 2017) discrepancies were observed between the perceptions and expectations of the protagonists in the interaction (Hart et al., 2006). As the results show, in general, there is a correlation between the scores of the independent observers and the patients after the intervention. However, there is no such correspondence between the perception of the patients and the health personnel (the students) during the simulation. Paradoxically, the general empathy measurement (Jefferson Empathy Scale) self-reported by the students correlated with the perceptions of the patients and observers in the experimental groups at both centres. In this case, such discrepancies could be due to the students' appreciation of his/her performance being influenced by social desirability and/or the belief that it will influence his/her grade. By contrast, it would also be influenced by the fact that being more aware of their deficiencies caused them to score lower. According to Williams and Stickley (2010), this research suggests that patients' perceptions of being understood do not correlate with the perceptions of observers and therapists.

The slight improvements within the Jefferson Empathy Scale scores (not in a significant way) after the intervention remind us that it is a self-reported measure of the general empathy level noted by the students, not a measure of the level that they show in the specific condition of simulation. It might be moderately improved only over the time; in fact, the scores also fluctuated slightly in the control group. On the other hand, the lack of correlation in the scores provided by the students in a previous general measurement (Jefferson Empathy Scale) and the specific one after the simulation (RES), for the majority of the groups can indicate a greater difference between the motivation to empathise, the moral dimension of empathy (Grau et al., 2016; Neumann et al., 2012; Sulzer et al., 2016), and their own beliefs on how much they have empathised in a concrete situation.

Another important aspect is the follow-up of the sample. Empathy may increase significantly immediately after an intervention, but there is limited data available regarding long-term efficacy (Kelm et al., 2014). Generally, empathy diminishes over time, as has been demonstrated (Neumann et al., 2012), although several authors recently identified the existence of variability of empathic response and showed

Table 4 (continued)

	Jefferson						Care						Self-esteem learning					
	Pre		Post		Follow-up		Pre		Post		Follow-up		Post		Follow-up		Follow-up	
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2
SELF	0.332*	0.272	0.056	0.224	0.042	0.377*	0.054	-0.014	-0.099	-0.179	0.000	0.153	-	-	0.697 [§]	0.818 [§]	0.697 [§]	0.818 [§]
Learning	0.233	0.121	0.164	0.270	0.328*	0.359*	0.260	0.096	-0.024	-0.236	0.009	0.267	0.296*	0.313*	0.225	0.111	0.225	0.206
	0.172	0.018	0.289*	0.241	0.301*	0.187	0.326*	-0.118	0.396*	0.357*	0.389*	0.110	0.296*	0.123	0.313*	0.123	0.795 [§]	0.513 [§]
	0.232	0.006	0.456 [§]	-0.062	0.390*	0.077	0.161	0.099	0.386*	0.352*	0.544*	0.391*	0.096	0.097	0.225	0.513 [§]	0.795 [§]	0.206

E1 = Experimental Group, Centre 1; E2 = Experimental Group, Centre 2; A = Pre-test evaluation; B = Post-test evaluation; C = Monitoring; RES = Reynolds Empathy Scale; JEF = Jefferson Scale; SELF = self-esteem; (*) = $p < 0.05$; (§) = $p < 0.01$.

the presence of different models (Díaz-Narváez et al., 2017). Few studies have performed a follow-up assessment of improvements in empathic capacity (Hart et al., 2006; Nosek et al., 2014; Yang et al., 2014), although only in some cases was the improvement statistically significant (Hart et al., 2006). In our case, the scores improved significantly in the follow-up with respect to the baseline in all of the measurements recorded. Comparing the follow-up scores to those obtained after the training, most of the scores improved, although there were significant differences only for some variables and/or groups. Despite this, there were statistically significant differences between the variables measured during the simulation between the baseline and follow-up.

That said, research such as Kiosses-Vassilios et al. (2016) has indicated that there is not certain knowledge about the improvements in empathic behaviour over the long-term. In the previously mentioned systematic revision, out of 17 controlled trials selected, the tracking did not exceed a year in any of the studies. Those with significant results of the intervention did not note any improvement in a subsequent stage when the participants were followed for a longer period.

4.1. Study strengths

The design used in this study (modified cohort-control); the detailed information about the training sessions; the use of multiple measurements that include of all of the agents involved; the realising an intervention in two different centres with the involvement of diverse teachers, and the use of objective measurements reported by independent observers, helped to compensate for the limitations of the other studies.

4.2. Limitations

A limitation of the present study is the use of a convenience sample rather than a representative sample, given that the students had to receive the same training for ethical reasons.

There is evidence that self-reported measures are frequently inaccurate in the evaluation of communication behaviours (Nosek et al., 2014). This study has attenuated these limitations with the inclusion of behaviour assessment by independent observers and the triangulation of the information. These strategies have provided reliability to the data.

On the other hand, the level of self-esteem at baseline was not registered, and due to the increase in said variable during the tracking respect of the post-intervention, it would have been interesting to see the effect after the intervention in relation to the initial self-esteem levels.

Although a random experimental study minimises the influence of possible confounding variables, it has not been considered, as other authors have affirmed (Lor et al., 2015; Ward, 2016). Other variables might influence empathy such as the empathy levels before starting college, the beginning of clinical work experience, and/or personal or familial backgrounds related to the sanitary system contact.

5. Conclusions

Experiential training development has shown positive results in nursing students, and the results are promising in terms of retaining the competencies that have been learned. As observed in the follow-up, the results were maintained over time, improving the participants' initial empathic capacity. It would be desirable to conduct a greater follow-up to assess if these scores are maintained for a longer time.

Training improves the cognitive, emotional, and relational dimensions of empathy but seems to have a slight impact on the moral dimension or the motivation to practise empathy. Regarding the future, this should be investigated and actions directed toward placing nursing

Table 5
Correlations centre 2.

	Carkhuft						Res						Jefferson										
	Pre		Post		Follow-up		Pre		Post		Follow-up		Pre		Follow-up		Post						
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2					
Carkhuft	A	-	-	0.163	-0.122	-0.215	-0.339	0.320	0.212	0.376*	-0.101	-0.329	0.083	B	0.163	-0.122	0.052	0.017	0.244	-0.065	-0.038	0.080	0.041
	B	-0.215	-	0.052	-	0.138	0.138	0.073	0.244	-0.038	-	-0.222	0.041	C	-0.339	0.320	0.017	0.244	-0.065	-0.038	0.080	0.041	
	C	0.052	0.138	-	0.138	-	0.138	0.073	0.244	-0.038	-	0.138	0.073	A	-0.339	0.320	0.017	0.244	-0.065	-0.038	0.080	0.041	
RES	A	0.212	0.244	0.017	0.244	-0.222	0.443*	-	0.443*	0.341	0.477*	0.344	0.009	B	0.212	0.244	0.017	0.244	-0.065	-0.038	0.080	0.041	
	B	-0.101	-0.376*	-0.065	-	0.138	0.138	0.073	0.244	-0.038	-	0.138	0.073	C	-0.101	-0.376*	-0.065	-	0.138	0.138	0.073	0.244	
	C	0.073	0.244	0.017	0.244	-0.222	0.443*	-	0.443*	0.341	0.477*	0.344	0.009	A	0.212	0.244	0.017	0.244	-0.065	-0.038	0.080	0.041	
JEF	A	-0.329	0.083	0.080	0.080	0.354	0.672 [§]	0.101	0.344	0.009	0.582 [§]	-	0.475*	B	-0.329	0.083	0.080	0.080	0.344	0.009	0.582 [§]	-	
	B	0.044	0.433*	0.547 [§]	0.145	0.177	0.297	0.685 [§]	0.344	0.545 [§]	0.098	0.219	0.409*	C	0.044	0.433*	0.547 [§]	0.145	0.177	0.297	0.685 [§]	0.344	
	C	0.346	0.159	0.159	0.159	-0.196	0.060	0.303	0.303	0.474*	0.079	0.295	0.475*	A	-0.329	0.083	0.080	0.080	0.344	0.009	0.582 [§]	-	
CARE	A	-0.041	0.159	0.236	0.318	0.431*	-0.192	0.303	-0.385*	-0.044	-0.313	-0.095	-0.077	B	-0.041	0.159	0.236	0.318	0.431*	-0.044	-0.313	-0.095	
	B	0.176	-0.225	0.743 [§]	0.568 [§]	0.075	-0.042	-0.029	0.210	0.217	-0.019	-0.084	0.145	C	0.176	-0.225	0.743 [§]	0.568 [§]	0.075	-0.042	-0.029	0.210	
	C	0.085	0.355	0.355	0.355	0.596 [§]	-0.209	0.091	-0.049	0.129	-0.084	0.195	0.175	A	0.085	0.355	0.355	0.355	0.596 [§]	-0.209	0.091	-0.049	0.129
SELF	B	0.335	0.444*	0.227	0.075	-0.162	-0.076	0.091	0.466*	0.129	0.428	0.072	0.175	C	0.335	0.444*	0.227	0.075	-0.162	-0.076	0.091	0.466*	
	C	0.086	0.336	0.213	0.074	0.336	0.336	0.470*	0.470*	0.451*	0.580 [§]	0.468*	0.020	A	0.086	0.336	0.213	0.074	0.336	0.336	0.470*	0.470*	
	Learning	0.221	0.336	0.485*	-0.209	-0.156	0.284	0.304	0.429*	0.429*	0.172	0.131	0.020	B	0.221	0.336	0.485*	-0.209	-0.156	0.284	0.304	0.429*	
C	0.263	0.232	0.232	0.232	0.471*	0.277	0.173	0.173	0.173	0.335	0.479*		C	0.263	0.232	0.232	0.232	0.471*	0.277	0.173	0.173		

	Jefferson						Care						Self-esteem learning									
	Post		Follow-up		Pre		Post		Follow-up		Pre		Post		Follow-up		Pre					
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2				
Carkhuft	0.044	0.433*	0.346	0.159	-0.041	0.176	0.176	0.176	0.085	0.335	0.444*	0.086	0.547 [§]	0.145	0.159	0.236	0.318	0.743 [§]	0.227	0.075	0.213	0.263
	0.177	0.177	-0.196	0.318	0.431*	0.075	0.075	0.075	0.596 [§]	-	0.485*	0.232	0.177	0.177	-0.196	0.318	0.431*	0.075	0.075	0.075	0.075	0.263
	0.297	0.436*	0.060	0.303	-0.192	-0.042	-0.042	-0.042	-0.209	-0.209	-0.076	0.020	0.297	0.436*	0.060	0.303	-0.192	-0.042	-0.042	-0.042	-0.042	0.263
RES	0.545 [§]	0.098	0.474*	-0.044	-0.385*	0.210	0.210	0.210	-0.049	0.217	0.466*	0.470*	0.545 [§]	0.145	0.159	0.236	0.318	0.743 [§]	0.227	0.075	0.213	0.263
	0.219	0.079	0.079	-0.044	-0.313	-0.019	-0.019	-0.019	-0.084	0.428	0.429*	0.173	0.219	0.079	0.079	-0.044	-0.313	-0.019	-0.019	-0.019	-0.019	0.263
	0.409*	0.475*	0.295	-0.077	-0.095	-0.084	-0.084	-0.084	0.168	0.428	0.429*	0.173	0.409*	0.475*	0.295	-0.077	-0.095	-0.084	-0.084	-0.084	-0.084	0.263
JEF	-	-	-	0.507*	0.103	-	-	-	-0.216	-	-	0.169	-	-	-	-	-	-	-	-	-	0.263
	0.507*	0.374*	-	-0.349	-0.349	-0.123	-0.123	-0.123	0.235	0.399	0.399	0.169	0.507*	0.374*	-	-0.349	-0.349	-0.123	-0.123	-0.123	-0.123	0.263
	0.103	0.374*	-0.349	-	-	0.367*	0.367*	0.367*	0.235	0.399	0.399	0.169	0.103	0.374*	-0.349	-0.349	-0.123	-0.123	-0.123	-0.123	-0.123	0.263
CARE	0.416*	0.168	-0.123	0.235	0.367*	-	-	0.401	0.401	0.159	0.159	0.168	0.416*	0.168	-0.123	0.235	0.367*	0.367*	0.367*	0.367*	0.367*	0.263
	0.069	0.168	-0.216	0.401	0.507*	0.401	0.401	0.401	0.401	0.159	0.159	0.168	0.069	0.168	-0.216	0.401	0.507*	0.507*	0.507*	0.507*	0.507*	0.263
	0.069	0.168	-0.216	0.401	0.507*	0.401	0.401	0.401	0.401	0.159	0.159	0.168	0.069	0.168	-0.216	0.401	0.507*	0.507*	0.507*	0.507*	0.507*	0.263

(continued on next page)

Table 5 (continued)

	Jefferson				Care				Self-esteem learning			
	Post		Follow-up		Pre		Post		Post		Follow-up	
	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2	E1	E2
SELF	0.485*	0.392*	0.328	0.239	-0.020	0.239	0.195	-0.009	0.427	0.327	0.419*	0.327
			0.351		-0.040		0.350		0.253		0.737 [§]	0.278
Learning	0.394	0.077	0.399	0.384*	0.159	0.384*	0.394	0.085	0.387		0.586*	0.431*
	0.169		0.253		0.016		0.038		0.513*		0.419*	0.327
											0.594*	0.431*

E1 = Experimental Group; Centre 1; E2 = Experimental Group; Centre 2; A = Pre-test evaluation; B = Post-test evaluation; C = Monitoring; RES = Reynolds Empathy Scale; JEF = Jefferson Scale; SELF = self-esteem (*) = $p < 0.05$; (§) = $p < 0.01$.

students in their patients' shoes in order to strengthen said motivation/dimension.

Contributions

The authors have all made substantial contributions to the following: (1) the conception and design of the study (PBS, MFG), the acquisition of the data (PBS, MFG, MDR, iCARE Team), and the analysis and interpretation of the data (PBS, MFG); (2) drafting the article and revising it critically for important intellectual content (PBS, MFG, MDR, iCARE Team); and (3) final approval of the version to be submitted (PBS, MFG, MDR).

Funding

This work was supported by the Departmental Collaboration Grant (Ministry of Education, Culture and Sports); the Teaching Innovation Unit at the University of Cadiz [Teaching Innovation Projects, number: Sol-201500054514-tra].

This work won first prize in the call for Teaching Innovation Projects at the University of Cadiz in 2016.

Declaration of Competing Interest

None.

Acknowledgements

The authors thank the University of Cadiz for funding support for this project. They also express appreciation to all of the students who took part in this study and the collaborating students from the Nursing and Physiotherapy Department (Buchra Abdel-Lathahal, Alba Correro-Bermejo, Patricia de Ángeles-Estévez, Sonia de la Torre-Moyano, Marta Sánchez-Lobato, Pablo Sergio Soler-Martins, and Rafael Valdivia-Diaz) for their participation and logistical support.

The iCARE team was composed of Concepción Carnicer-Fuentes, Cristina Castro-Yuste, María José García-Cabanillas, Cristina Gavira-Fernández, María de los Ángeles Martelo-Baro, Olga Paloma-Castro, María del Carmen Paublete-Herrera, María Jesús Rodríguez-Cornejo, and Luis Moreno-Corral. They assisted in the evaluation.

The authors would like to thank the collaboration of Jose Manuel Romero-Sánchez for his assistance in the randomised distribution.

References

Bas-Sarmiento, P., Fernández-Gutiérrez, M., Baena-Baños, M., Romero-Sánchez, J.M., 2017. Efficacy of empathy training in nursing students: a quasi-experimental study. *Nurse Educ. Today* 59, 59–65. <https://doi.org/10.1016/j.nedt.2017.08.012>.

Batt-Rawden, S.A., Chisolm, M.S., Anton, B., Flickinger, T.E., 2013. Teaching empathy to medical students: an updated, systematic review. *Acad. Med.* 88, 1171–1177. <https://doi.org/10.1097/ACM.0b013e318299f3e3>.

Bellver-Navalón, C., Marco-Gisbert, A., Reynolds, W., Muñoz-Izquierdo, A., Fonfria-Vivas, C., Fonfria-Vivas, R., Fusté-Rodríguez, D., Zubicoa-Fernández, C., 2004. La importancia de una buena traducción. *Rev ROL Enferm.* 27, 857–862.

Bikker, A.P., Fitzpatrick, B., Murphy, D., Mercer, S.W., 2015. Measuring empathic, person-centred communication in primary care nurses: validity and reliability of the consultation and relational empathy (CARE) measure. *BMC Fam. Pract.* 16, 149. <https://doi.org/10.1186/s12875-015-0374-y>.

Carkhuff, R.R., 1969. *Helping and human relations*. In: *Practice and research*. 11. Holt, Rinehart & Winston, New York, pp. 319–320.

Chen, I., Forbes, C., 2014. Reflective writing and its impact on empathy in medical education: systematic review. *J. Educ. Eval. Health Prof.* 11, 20. <https://doi.org/10.3352/jeehp.2014.11.20>.

Díaz, V.P., Alonso, L.M., Caro, S.E., Silva, M.G., Arboleda, J., Bilbao, J.L., Iglesias, J., 2014. Orientación empática de estudiantes de medicina en tres universidades de Barranquilla, Colombia y en una universidad de República Dominicana. *Arch. Argent. Pediatr.* 112, 41–51. <https://doi.org/10.5546/aap.2014.e41>.

Díaz-Narváez, V.P., Coronado, A.M., Bilbao, J.L., González, F., Padilla, M., Calzadilla-Núñez, A., Silva-Vetri, M.G., Arboleda, J., Bullen, M., Utsman, R., Fajardo, E., Alonso, L.M., Cervantes, M., Varela, T., 2017. Reconsidering the 'decline' of dental student empathy within the course in Latin America. *Acta Medica Port.* 30 (11), 775–782. <https://doi.org/10.20344/amp.8681>.

- Grau, A., Toran, P., Zamora, A., Quesada, M., Carrion, C., Vilert, E., Castro, A., Cerezo, C., Vargas, S., Gali, B., Cordon, F., 2016. Evaluación de la empatía en estudiantes de Medicina. *Educ. Med.* 18 (2), 114–120. <https://doi.org/10.1016/j.edumed.2016.04.001>.
- Hart, C.N., Drotar, D., Gori, A., Lewin, L., 2006. Enhancing parent-provider communication in ambulatory pediatric practice. *Patient Educ. Couns.* 63 (1–2), 38–46. <https://doi.org/10.1016/j.pec.2005.08.007>.
- Hojat, M., Mangione, S., Nasca, T.J., Cohen, M.J.M., Gonnella, J.S., Erdmann, J.B., Veloski, J., Magee, M., 2001. The Jefferson scale of physician empathy: development and preliminary psychometric data. *Educ. Psychol. Meas.* 61, 349–365. <https://doi.org/10.1177/00131640121971158>.
- İlhan, N., Sukut, O., Utaş Akhan, L., Batmaz, M., 2016. The effect of nurse education on the self-esteem and assertiveness of nursing students: a four-year longitudinal study. *Nurse Educ. Today* 39, 72–78. <https://doi.org/10.1016/j.nedt.2015.12.026>.
- Kelm, Z., Womer, J., Walter, J.K., Feudtner, C., 2014. Interventions to cultivate physician empathy: a systematic review. *BMC Med. Educ.* 14, 219. <https://doi.org/10.1186/1472-6920-14-219>.
- Kiosses-Vassilios, N., Karathanos-Vassilios, T., Tatsioni, A., 2016. Empathy promoting interventions for health professionals: a systematic review of RCTs. *J. Compassionate Health Care* 1–22. <https://doi.org/10.1186/s40639-016-0024-9>.
- Lobchuk, M., Halas, G., West, C., Harder, N., Tursunova, Z., Ramraj, C., 2016. Development of a novel empathy-related video-feedback intervention to improve empathic accuracy of nursing students: a pilot study. *Nurse Educ. Today* 46, 86–93. <https://doi.org/10.1016/j.nedt.2016.08.034>.
- Lor, K.B., Truong, J.T., Ip, E.J., Barnett, M.J., 2015. A randomized prospective study on outcomes of an empathy intervention among second-year student pharmacists. *Am. J. Pharm. Educ.* 79 (2), 1–4. <https://doi.org/10.5688/ajpe79218>.
- Martín-Albo, J., Núñez, J.L., Navarro, J.G., Grijalvo, F., 2007. The Rosenberg self-esteem scale: translation and validation in university students. *Span. J. Psychol.* 10 (2), 458–467.
- McKillip, J., 1979. Impact evaluation of service programs: three flexible designs. *Evaluation Quart.* 3, 97–104.
- Mercer, S.W., Reynolds, W.J., 2002. Empathy and quality of care. *Br. J. Gen. Pract. Supp.* 52, S9–S12.
- Mercer, S.W., Maxwell, M., Heaney, D., Watt, G.C., 2004. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam. Pract.* 21, 699–705. <https://doi.org/10.1093/fampra/cmh621>.
- Neumann, M., Scheffer, C., Tauschel, D., Lutz, G., Wirtz, M., Edelhäuser, F., 2012. Physician empathy: definition, outcome-relevance and its measurement in patient care and medical education. *GMS Z. Med. Ausbildung.* 29 (1), Doc11. <https://doi.org/10.3205/zma000781>.
- Nosek, M., Gifford, E., Kober, B., 2014. Nonviolent communication (NVC) training increases empathy in baccalaureate nursing students: a mixed method study. *J. Nurs. Educ. Pract.* 4, 1–16. <https://doi.org/10.5430/jnep.v4n10p1>.
- Pasli-Gurdogan, E., Çentinkaya-Uslusoy, E., Kurt, S., Yasak, K., 2016. Comparison of the self-esteem and communication skills at the 1st and senior year nursing students. *IJCS* 9 (2), 496–502.
- Reynolds, W., 2000. *The Measurement and Development of Empathy in Nursing*. Ashgate, Aldershot.
- Sulzer, S.H., Feinstein, N.W., Wendland, C.L., 2016. Assessing empathy development in medical education: a systematic review. *Med. Educ.* 50 (3), 300–310. <https://doi.org/10.1111/medu.12806>.
- Ward, J., 2016. The empathy enigma: does it still exist? Comparison of empathy using students and standardized actors. *Nurse Educ.* 41 (3), 134–138. <https://doi.org/10.1097/NNE.000000000000236.0>.
- Ward, J., Cody, J., Schaal, M., Hojat, M., 2012. The empathy enigma: An empirical study of decline in empathy among undergraduate nursing students. *J. Prof. Nurs.* 28, 34–40. <https://doi.org/10.1016/j.profnurs.2011.10.007>.
- Williams, B., Boyle, M., Fielder, C., 2015. Empathetic attitudes of undergraduate paramedic and nursing students towards four medical conditions: a three-year longitudinal study. *Nurse Educ. Today* 35 (2), e14–e18. <https://doi.org/10.1016/j.nedt.2014.12.007>.
- Williams, J., Stickley, T., 2010. Empathy and nurse education. *Nurse Educ. Today* 30, 752–755. <https://doi.org/10.1016/j.nedt.2010.01.018>.
- Yang, K., Woomey, G.R., Agbemenu, K., Williams, L., 2014. Relate better and judge less: poverty simulation promoting culturally competent care in community health nursing. *Nurse Educ. Pract.* 14 (6), 680–685. <https://doi.org/10.1016/j.nepr.2014.09.001>.