Original Article/*Özgün Araştırma*
doi: 10.54875/jarss.2024.57873

Stereotype Threat Effect on Cardiopulmonary Resuscitation: A Randomized Controlled Mannequin Study*

Kalıpyargı Tehditinin Kardiopulmoner Resüsitasyon Performansı Üzerine Etkisi: Bir Randomize Kontrollü Manken Çalışması

Murat Tumer[1], Leman Korkmaz[2], Filiz Uzumcugil[3], Aysun Ankay Yilbas[3], Banu Kilicaslan[3], Seda Banu Akinci[3]

[1]VKV Amerikan Hospital, Clinic of Anesthesiology and Reanimation, İstanbul, Türkiye
[2]Middle East Technical University, Faculty of Arts and Science, Department of Psychology, Ankara, Türkiye
[3]Hacettepe University Faculty of Medicine, Department of Anesthesiology and Reanimation, Ankara, Türkiye

*The first draft of this article was published as a preprint on the website https://www.researchsquare.com/article/rs-3196901/latest

ABSTRACT

Objective: Stereotype threat (ST) can lead to decreased performance when individuals face the possibility of confirming negative stereotypes associated with their group. During the Coronavirus disease 2019 (COVID-19) pandemic, non-Intensive Care Unit physicians (non-ICUp) were assigned to work in ICUs. However, social media emphasized the inadequacy of knowledge and skills among these physicians. Given the negative judgments, the study aimed to evaluate the cardiopulmonary resuscitation (CPR) performances of these physicians and investigate the effect of ST.

Methods: A total of 63 non-ICUp and 53 Intensive Care Unit physicians (ICUp) physicians working in COVID-19 ICUs were randomly assigned to control and experimental groups. In the experimental group, ST was manipulated by presenting the study's aim as measuring the difference in CPR performances between ICUp and non-ICUp physicians. The control group received no information. Participants were videotaped while performing a standard CPR scenario and evaluated by independent instructors and mannequin scores.

Results: Overall CPR scores were higher among ICUp. Non-ICUp performed better in the ST condition regarding effective chest compression (p=.02) and correct compression rates per minute (p=.02) compared to the control condition. However, ICUp had lower scores for correctly placing chest compressions in the ST condition (p=.03).

Conclusion: The higher CPR performance among ICUp was expected. However, the hypothesis suggesting lower performance for non-ICUp under ST conditions was not supported. Inconsistent results regarding the ST effect could be influenced by moderating factors such as task difficulty, knowledge about the existing stereotype, and motivation to perform well. The interaction between the physicians' specialty and situational factors highlights the importance of creating realistic training environments that

ÖZ

Amaç: Kalıpyargı tehditi (KT), ait oldukları gruplarla ilişkilendirilen olumsuz kalıpyargılar nedeniyle insanların performanslarında düşüşe yol açabilir. Covid-19 pandemisi sırasında, yoğun bakım ünitesinde çalışmayan doktorlar (non-YBU) yoğun bakım birimlerinde görevlendirildiler. Ancak sosyal medyada bu doktorların bilgi ve becerilerindeki yetersizlik vurgulandı. Çalışmamız, negatif kalıpyargılar göz önüne alındığında, yoğun bakımlarda çalışan doktorların kardiyopulmoner resüsitasyon (KPR) performanslarını değerlendirmeyi ve KT'nin doktorların performansları üzerine etkisini araştırmayı amaçlamıştır.

Yöntem: Toplamda 63 non-YBU ve 53 COVID-19 yoğun bakım ünitesi (YBÜ)'de çalışan YBU doktoru, kontrol ve deney gruplarına rastgele atandı. Deney gruplarına çalışmanın amacının non-YBU ve YBU doktorlarının KPR performansları arasındaki farkı ölçmek olduğu söylendi. Böylece bu gruplara kalıpyargı tehditi verilmiş oldu. Kontrol grubuna bu bilgi verilmedi. Katılımcılar, standart bir KPR senaryosunu gerçekleştirirken videoya çekildi ve bu videolar bağımsız eğitmenler ve manken puanları ile değerlendirildi.

Bulgular: Genel KPR puanları YBU doktorlarında daha yüksekti. Non-YBU doktorları, etkili göğüs kompresyonu (p=.02) ve dakikadaki kompresyon sayıları (p=.02) açısından kontrol gruplarına göre KT koşulunda daha iyi performans sergiledi. Ancak YBU doktorlarında KT koşulunda doğru noktaya göğüs kompresyonu yapma puanları daha düşüktü (p=.03).

Sonuç: Yoğun bakım ünitesinde çalışan doktorlarda daha yüksek KPR performansı beklenen bir bulguydu. Ancak KT koşullarında non-YBU doktorlarındaki düşük performans hipotezimiz desteklenmedi. Kalıpyargı tehdidi etkisiyle ilgili tutarsız sonuçlar, görevin zorluğu, mevcut kalıpyargı hakkındaki bilgi ve iyi performans gösterme motivasyonu gibi düzenleyici faktörlerden etkilenmiş olabilir. Doktorların uzmanlığı ve durumsal faktörler arasındaki etkileşim, yüksek baskı durumlarını simüle eden gerçekçi eğitim

Received/Geliş tarihi : 27.12.2023
Accepted/Kabul tarihi : 09.01.2024
Publication date : 31.01.2024

***Corresponding author:** Murat Tumer • dr.m.tumer@gmail.com
Murat Tumer 0000-0001-9132-9992 / Leman Korkmaz 0000-0003-2755-7290
Filiz Uzumcugil 0000-0001-9161-3248 / Aysun Ankay Yilbas 0000-0001-6196-3191
Banu Kilicaslan 0000-0003-3295-9999 / Seda Banu Akinci 0000-0002-8538-6955

Cite as: Tumer M, Korkmaz L, Uzumcugil F, Ankay Yilbas A, Kilicaslan B, Akinci SB. Stereotype threat effect on cardiopulmonary resuscitation: A randomized controlled mannequin study. JARSS 2024;32(1):36-45.

simulate high-pressure situations, ultimately contributing to the development of competent and confident healthcare professionals. Future research should further explore the impact of ST-based training on interactions and performance among different healthcare professionals.

Keywords: Stereotype threat, intensive care unit, cardiopulmonary resuscitation, COVID-19

ortamlarının oluşturulmasının önemini vurgulamıştır. Bu eğitim proglamları yetkin ve kendine güvenen sağlık profesyonellerinin gelişimine katkıda bulunabilir. Gelecekteki araştırmalar, KT temelli eğitimin farklı sağlık profesyonelleri arasındaki etkileşimi ve performansları üzerindeki etkisini daha fazla araştırmalıdır.

Anahtar sözcükler: Kalıpyargı tehditi, yoğun bakım ünitesi, kardiyopulmoner resüsitasyon, Covid-19

INTRODUCTION

Stereotypes are beliefs about the characteristics, behaviors, and attributes of different groups (1). Stereotype threat (ST) is a psychological discomfort experienced by individuals when they have the possibility of confirming negative stereotypes about their social group, leading to decreased performance (2). Several factors such as anxiety, low-performance expectations, physiologic arousal, and reduced memory capacity can contribute to ST (3). Schmader et al. proposed a model explaining the mechanism of ST, focusing on stress stimulation, performance monitoring, and efforts to suppress thoughts and emotions (4).

Most research on ST has been conducted in laboratory settings (5). However, to ensure the generalizability of the results, studies conducted in operational settings are needed (5). In this study, we examined beliefs and stereotypes about the competence of non-intensive care unit physicians (non-ICUp) working in ICUs during the coronavirus disease 2019 (COVID-19) pandemic, collecting data in a hospital setting.

Due to the COVID-19 pandemic, there has been a significant increase in patients in ICUs worldwide since 2020 (6). As a result, the number of physicians in ICUs has been insufficient, and non-ICUp from various medical specialties have been assigned to work in ICUs. Negative stereotypes have emerged, suggesting that non-ICUp lack sufficient knowledge and skills in intubation and treating lung infections (7,8). Despite the existence of these stereotypes, no studies have examined the differences in clinical performance between intensive care unit physicians (ICUp) and non-ICUp and how stereotypes affect the performance of non-ICUp in the ICU.

This study aimed to test the effect of ST on non-ICUp by evaluating the performance of ICUp and non-ICUp in cardiopulmonary resuscitation (CPR), which is a fundamental skill expected from all physicians working in the ICU, and ICUp are typically more experienced in its performance. We hypothesized that non-ICUp would have lower CPR performance compared with ICUp, and under ST conditions, non-ICUp would receive lower scores on CPR performance compared with non-ICUp without ST. We expected no difference in the performance of ICUp under ST and non-ST conditions.

MATERIAL and METHODS

The study was approved by the Hacettepe University Non-invasive Clinical Research Ethics Board (07.09.2021, GO21/902). The clinical trials registration code is NCT05074446.

Participants

The participants consisted of resident physicians from Hacettepe University who worked in the COVID-19 ICU. Anesthesiology residents who worked in the ICU for more than 3 months and had completed at least 1 year of their residency were classified as "ICUp." Residents who were assigned to work in the ICU due to the pandemic but would not have worked there under normal circumstances were classified as "non-ICUp." The non-ICUp participants came from various medical specialties (e.g., otorhinolaryngology, orthopedics, psychiatry, urology, and ophthalmology). The assignments were organized by the hospital management, and one resident from each specialty worked together with ICUp in the COVID-19 ICU according to a predetermined schedule. All non-ICUp participants received training on COVID precautions and basic care for patients with COVID-19. To attain the final number, invitations were extended to all research associates actively engaged in COVID-19 ICUs at our institution.

Materials and Procedure

Randomization was conducted using a stratified approach, wherein participants were initially grouped based on their medical specialty (ICUp and non-ICUp). Within each specialty subgroup, random assignment was performed using a computer program to ensure an equitable representation of individuals in both the experimental (ST) and control (no-threat) conditions. Considering the stereotype threat studies, a triple classification can be made for the manipulations used as (a) explicit, (b) partially explicit, and (c) indirect/implicit (9). When it is emphasized that one group is inferior or unsuccessful compared to the other group (e.g. men outperform women on this math test) it is possible to say that this is overt stereotyping manipulation. If the participants are told that there are group differences in the task without specifying the direction of the group difference (e.g. men and women perform differently on this math test), it is possible to classify this as partially explicit stereotyping manipulation. Indirect/

implicit stereotyping manipulations emphasize the diagnostic nature of the test (e.g. it is an intelligence test) or ask the participant's group identity before performing. We thought that the partially explicit method was appropriate in our study. The ST manipulation was performed on the experimental group just before they entered the testing room by a researcher who did not participate in the evaluation of the test. The participants were informed that the CPR performances of ICUp and non-ICUp working in the COVID-19 ICU would be compared. The control group was taken to the testing room without this explanation.

Participants were video recorded while performing a standard CPR scenario using a simulator mannequin (Resusci Anne Simulator®, 'Laerdal Medical') under the supervision of a researcher who was blinded to group allocation of the participants (Figure 1). The video recordings were scored separately by three independent internationally certified CPR expert instructors who were blinded to the participants' medical specialties and assigned groups. The experts evaluated the participants' performance using a 4-point scale on six different dimensions (Appendix 1). The performance was also measured using the simulator mannequin, with scores on three different dimensions using a 4-point scale (Appendix 2). After the test, participants were given a demographic form with questions about age, sex, medical specialty, employment duration, previous CPR experience, and CPR certification. A debriefing was provided at the end of the experiment.

Figure 1: CPR simulation scenario.

Design

The experiment used a 2X2 between-subjects design to predict CPR performance as evaluated by experts and the simulator mannequin. The independent variables were the physicians' medical specialty (ICUp and non-ICUp) and the ST condition (ST and no-threat conditions). Four groups were compared: ICUp in the experimental group, ICUp in the control group, non-ICUp in the experimental group, and non-ICUp in the control group.

Statistical Analysis

The statistical analysis were conducted using the Statistical Package for the Social Sciences (SPSS) Ver 28. The assumptions of the analysis, such as normality, homogeneity of variance, linearity, and homogeneity of regression slopes, were tested. The inter-rater reliability for the expert scores was calculated using Krippendorff's Alpha test (10). The interrater agreements among the three raters for the six different criteria ranged between 0.832 and 0.703, with scores above the lowest convincible limit (0.667) (11).

RESULTS

A total of 86 non-ICUp associates initially agreed to participate in the study; however, considering the working conditions during the pandemic and the operational arrangements within their respective departments, the research was conducted with a final cohort of 63 non-ICUp. A total of 59 ICUp agreed to participate in the study; however, due to similar reasons, the study was conducted with a cohort of 53 ICUp (Figure 2). The participants' ages ranged from 24 to 37 years (M=28.37, SD=2.32). Among the participants, 46% (n=53; 33 women and 20 men) were non-ICUp, and 54% (n=63; 20 women and 43 men) were ICUp. The demographic characteristics of the randomly assigned participants in each group are presented in Table I.

Independent samples t-tests showed that ICUp scored higher than non-ICUp on the total number of CPRs performed, the number of CPRs performed in the ICU, and the experience/duration in the ICU (Table II). A Chi-square test of independence indicated that 47 out of 53 ICUp had a CPR certificate, and only two out of 63 non-ICUp had a certificate (χ^2 (1, n=116) = 85.26, the phi coefficient -0.86).

Spearman's rank correlation was computed to assess the relationship between the study variables and demographic factors (Table III). Sex was significantly correlated with the observational expert score OB5, OB6, mannequin score M2, and M3. Age and employment duration did not correlate with the expert scores or the simulator mannequin scores.

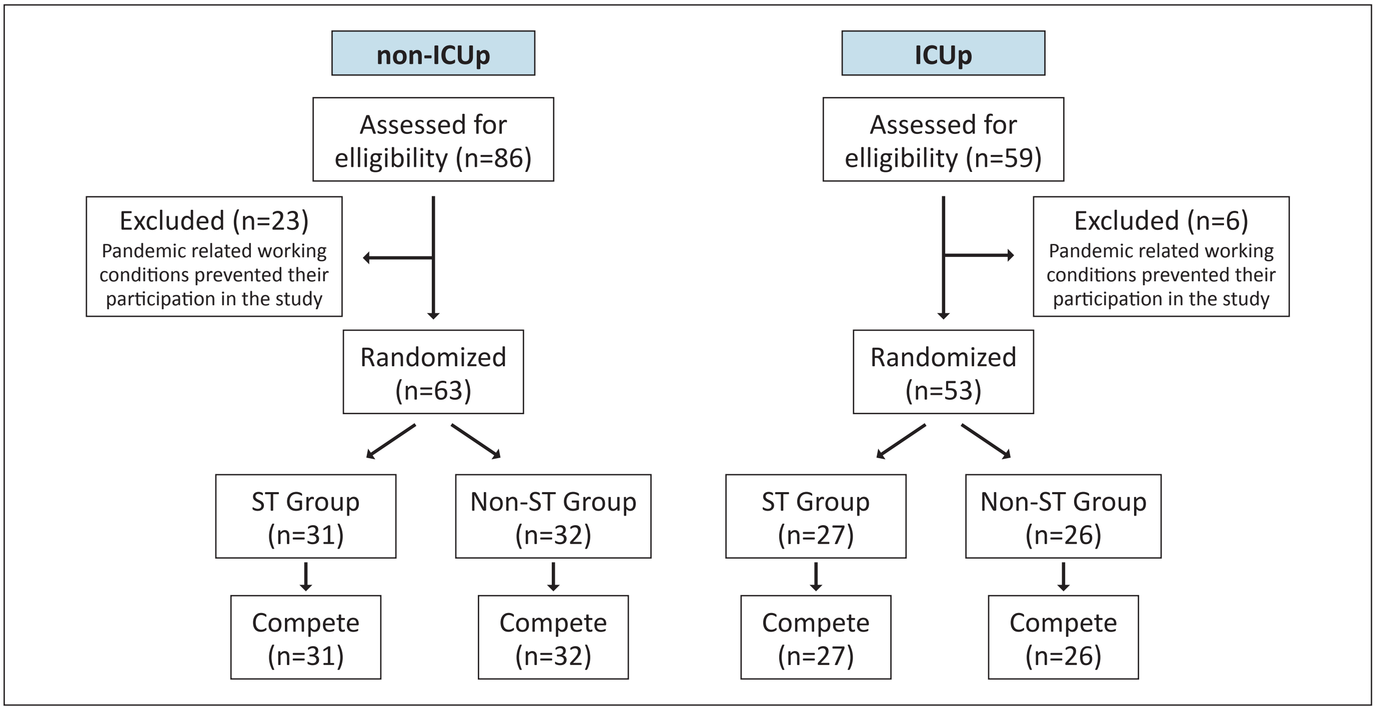

Figure 2: Participant flow diagram.

Table I: Demographic Characteristics of the ICUp and non-ICUp in the Threat/ No Threat Conditions

			Gender	Age		Term of Employment	
		n	Frequency (W/M)	M	SD	M	SD
ICUp	Control	26	16/10	29.00	2.25	31.00	16.42
	Experimental	27	17/10	28.70	2.85	31.50	17.35
Non-ICUp	Control	32	9/22	28.50	1.88	37.88	15.39
	Experimental	31	11/21	27.42	2.09	27.61	18.69

Term of employment was measured as a month. **W:** Women, **M:** Men, **ICU:** Intensive care unit, **ICUp:** ICU physicians, **Non-ICUp:** Non-ICU physicians.

Table II: Group Comparisons for ICUp and non-ICUp for Previous CPR Experiences, and Experience in ICU

	ICUp		Non-ICUp			
	M	SD	M	SD	t(114)	p
Total CPR experiences	23.91	25.26	12.60	19.43	2.66	.009
CPR experiences in the ICU	9.77	9.43	1.19	1.90	6.52	< .001
Experience in ICU	139.81	70.94	36.52	17.71	10.33	< .001

ICUp: ICU physicians, **Non-ICUp:** non-ICU physicians, **t:** Experience in ICU was measured as a day, **CPR:** Cardiopulmonary Resuscitation.

A 2X2 between-subjects analysis of variance (ANOVA) was conducted to analyze the interaction of medical specialty and ST condition in predicting performances on OB1, OB2, OB3, OB4, and M1. Sex was included as a covariate in the analysis based on the significant correlations with OB5, OB6, M2, and M3 (Table III). Post hoc comparisons using Bonferroni correction were calculated for significant interactions. The means, standard deviations, and significant mean differences for ICUp and non-ICUp in the threat/no-threat conditions are presented in Table IV.

The ANOVA results showed a significant main effect of medical specialty on OB1, OB2, OB3, OB4, and a marginal effect on M1. Intensive care unit physicians performed better than non-ICUp on these measures. The main effect of the ST condi-

Table III: Correlations, Means, and Standard Deviations of Demographic Factors and Study Variables

	Gender	Age	Toe	OB1	OB2	OB3	OB4	OB5	OB6	M1	M2	M3
Gender	-	.092	.118	.004	.156	.176	.131	.315**	-.228*	.043	.530**	-.258**
Age		-	.737**	-.012	-.024	.103	.035	-.004	.051	.134	.004	.139
Term of Employment			-	.048	-.077	.078	.085	-.016	-.058	.137	.062	.111
OB1				-	.674**	.560**	.461**	.467**	.394**	.347**	.279**	.317**
OB2					-	.544**	.425**	.450**	.245**	.261**	.423**	.132
OB3						-	.735**	.835**	.554**	.427**	.481**	-.030
OB4							-	.558**	.385**	.587**	.364**	.015
OB5								-	.488**	.255**	.657**	-.107
OB6									-	.360**	.035	.170
M1										-	.096	.028
M2											-	-.217*
M3												-
	M	28.37	32.11	3.14	3.03	2.43	2.50	2.79	2.91	2.78	2.91	2.48
	SD	2.32	17.23	1.01	0.90	1.13	1.14	0.97	0.86	1.26	1.26	1.32

****:** Correlation is significant at the .01 level (2-tailed), ***:** Correlation is significant at the .05 level (2-tailed). **Gender was dummy coded: 0**: Women and **1:** Men; **Toe:** Term of Employment was measured in a month, **OB1:** Observational Expert Score 1 (Are the chest compressions applied to the correct place?), **OB2:** Observational Expert Score 2 (Are the hands placed correctly?), **OB3:** Observational Expert Score 3 (Is effective compression applied?), **OB4:** Observational Expert Score 4 (Is the appropriate compression number per minute applied?), **OB5:** Observational Expert Score 5 (Is the compression depth appropriate?), **OB6:** Observational Expert Score 6 (Is the rib cage allowed to retract?), **M1:** Mannequin Score 1 (Chest compression number per minute), **M2:** Mannequin Score 2 (Chest compression depth), **M3:** Mannequin Score 3 (Are the chest compressions applied to the correct place?).

Table IV: Means and Standard Deviations and Significance and Confidence Intervals of Mean Differences for ICUp and non-ICUp in the Threat/No Threat Conditions

		OB1		OB2		OB3		OB4		OB5		OB6		M1		M2		M3	
		M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
ICUp	**No-Threat**	3.67	0.49	3.39	0.70	2.97	.99	2.78	1.11	3.15	0.89	3.52	0.35	3.00	1.67	2.74	1.36	2.99	1.26
	Threat	3.07	1.03	3.14	0.82	2.73	1.11	2.80	1.11	3.12	0.87	3.33	0.76	3.00	1.14	3.02	1.33	2.59	1.22
	Significance of Difference	p = .03		p = .31		p = .39		p = .95		p = .91		p = .36		p = 1.00		p = .35		p = .13	
	CI for Difference	[0.58, 1.13]		[-0.23, 0.73]		[-0.32, 0.82]		[-0.62, 0.58]		[-0.48, 0.56]		[-0.21, 0.57]		[-0.67, 0.67]		[-0.78, 0.41]		[-0.15, 1.21]	
Non-ICUp	**No-Threat**	2.83	1.12	2.76	0.88	1.78	.90	1.96	1.03	2.35	1.05	2.40	0.77	2.22	1.34	2.92	1.17	2.47	1.32
	Threat	3.13	1.04	2.94	1.05	2.41	1.16	2.57	1.17	2.67	0.96	2.56	0.83	2.97	1.22	2.92	1.15	1.96	1.26
	Significance of Difference	p = .23		p = .43		p = .019		p = .03		p = .16		p = .39		p = .02		p = .99		p = .05	
	CI for Difference	[-0.78, 0.19]		[-0.62, 0.27]		[-1.15, -0.11]		[-1.16, -0.06]		[-0.83, 0.12]		[-0.52, 0.20]		[-1.36, -0.14]		[-0.45, 0.64]		[0.01, 1.26]	

p-values were given for the significance of the mean differences based on comparison of ICUp in the threat and no-threat conditions and the comparisons of non-ICUp in the threat and no-threat conditions. For OB5, OB6, M2 and M3, adjusted means controlling for gender were given. CI for Difference represents 95% confidence intervals for mean differences. **OB1:** Observational Expert Score 1 (Are the chest compressions applied to the correct place?), **OB2:** Observational Expert Score 2 (Are the hands placed correctly?), **OB3:** Observational Expert Score 3 (Is effective compression applied?), **OB4:** Observational Expert Score 4 (Is the appropriate compression number per minute applied?), **OB5:** Observational Expert Score 5 (Is the compression depth appropriate?), **OB6:** Observational Expert Score 6 (Is the rib cage allowed to retract?), **M1:** Mannequin Score 1 (Chest compression number per minute), **M2:** Mannequin Score 2 (Chest compression depth), **M3:** Mannequin Score 3 (Are the chest compressions applied to the correct place?), **M:** Mean, **SD:** Standard deviation.

tion was not significant for OB1, OB2, OB3, and OB4. However, for M1, there was marginal significance, with performance being higher in the threat condition compared with the no-threat condition. The interaction between medical specialty and ST condition was significant for OB1 and OB3, and marginally significant for M1. Post hoc comparisons revealed that ICUp performed better on OB1 in the no-threat condition compared with the threat condition. There was no significant difference in the performance of non-ICUp between the no-threat and threat conditions (Figure 3). For OB3, non-ICUp performed better in the threat condition compared with the no-threat condition. There was no significant difference in the performance of ICUp between the no-threat and threat conditions (Figure 4). In predicting M1, ICUp showed no significant difference in performance between the threat and no-threat conditions, whereas non-ICUp performed better in the threat condition (Figure 5).

In predicting OB5, OB6, M2, and M3, a 2X2 between-subjects analysis of covariance (ANCOVA) was conducted, including sex as a covariate. Sex significantly predicted OB5, M2, and M3, but not OB6. The main effect of medical specialty was significant for OB5, OB6, and marginally significant for M3, with ICUp performing better than non-ICUp. The main effect of the ST condition was not significant for OB5, OB6, M2, and M3. The interaction between medical specialty and ST condition was not significant for any of these measures.

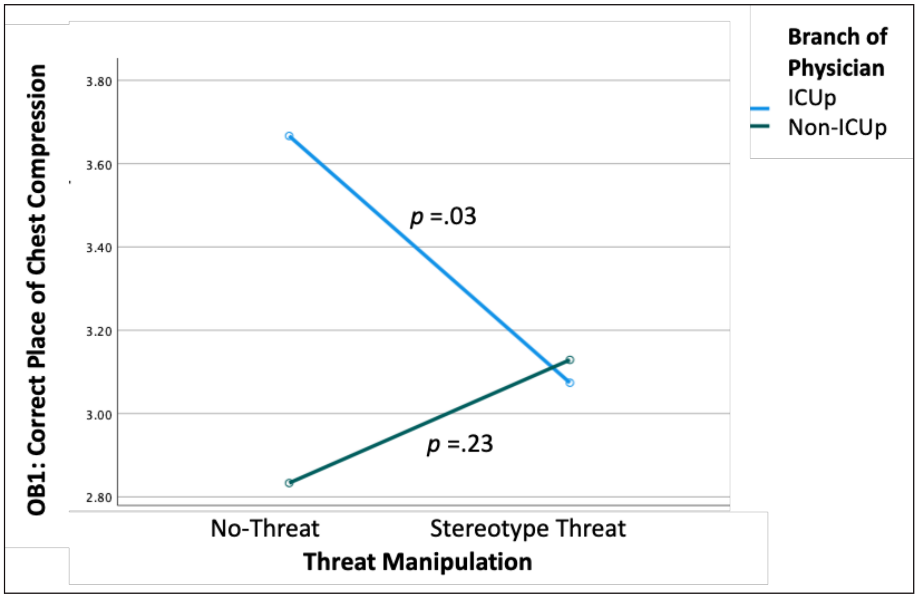

Figure 3: The interaction of branch of physicians and ST manipulation in predicting the correct place of chest compression score.

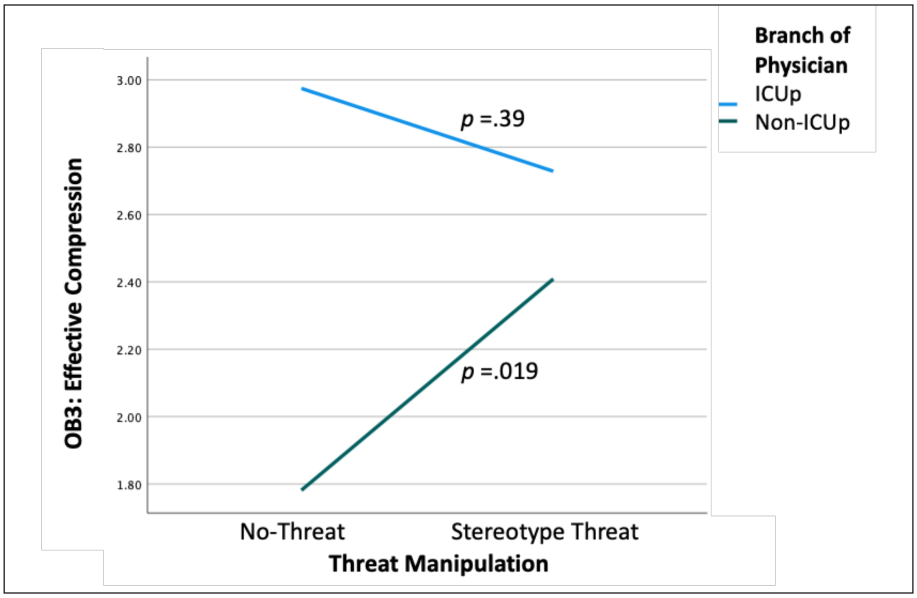

Figure 4: The interaction of branch of physicians and ST manipulation in predicting the effective compression score.

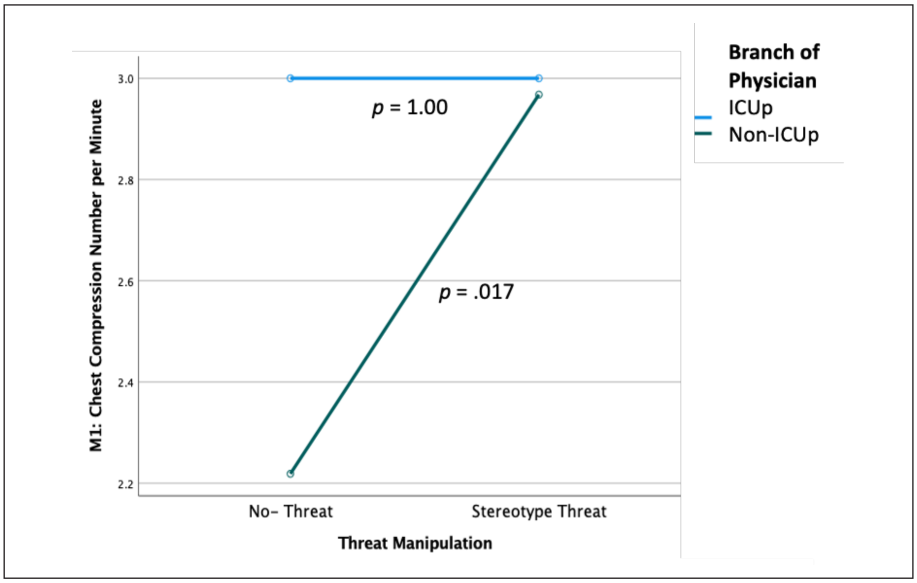

Figure 5: The interaction of branch of physicians and ST manipulation in predicting the chest compression number score.

Analysis Assumptions

For A 2X2 between-subjects ANOVA models, which were used in predicting OB1, OB2, OB3, OB4, and M1, outliers within each group were detected and assumptions of normality indicating normal distribution of means within each group and homogeneity of variance were tested. For the A 2X2 between-subjects ANCOVA models, which were used in predicting OB5, OB6, M2, and M3, in addition to the assumptions of between-subjects ANOVA assumptions, the assumption suggesting that the relationship between each covariate (CV) and the dependent variable (DV) were linear and homogeneity of regression slopes assumption were tested.

For the normality assumption, skewness, and kurtosis values (+/- 3 were accepted as cut-offs for these values) were calculated and quantile-quantile (Q-Q) plots for each cell/group were created to examine the shape of the distributions. Except for OB1, skewness and kurtosis values were in an acceptable range (scores ranged between -1.67 and 1.96). For OB1, the kurtosis value for the ICUp in the control condition was 4.24; one participant featured as an outlier. After excluding the outlier from the data, the kurtosis score obtained was 1.00. Thus, further analysis for OB1 was conducted without the outlier. Q-Q plots created for each group on all dependent variables lied approximately on a straight line, which indicated a normal distribution.

For testing the homogeneity of variance, Levene's test of homogeneity of variance (p-values greater than .05 were accepted as an indication of homogeneity of variance) and F*max* representing the ratio of the largest cell variance to the smallest one (10 is acceptable for approximately equal cell size) were used. Levene's test of homogeneity of variance revealed that the homogeneity of variance assumption was met by all variables except OB1 and OB6. However, the F*max* scores for OB1 (5.29) and OB6 (5.5) indicated acceptable scores for homogeneity of variance.

For between-subjects ANCOVA models used to predict OB5, OB6, M2, and M3, the linearity of the relationship between each CV and the DV were tested using Spearman's rank correlation. In the current study, variables showing a significant linear correlation with dependent variables were included in further analysis as covariates. Homogeneity of regression slopes assumption suggests that the slopes of regression of the DV on the covariates do not differ across groups. This assumption was tested by calculating interactions between the independent variable and covariate in predicting OB5, OB6, M2, and M3. The interactions between the independent variables and covariates (sex was dummy coded) were not significant, thus for all variables, assumptions of homogeneity of regression were satisfactory.

DISCUSSION

Stereotype effects on minority groups, such as those based on sexual and ethnic identity, and on health workers due to their sex, have been studied in the literature (12-14). However, there is limited research on how stereotypes impact performance. Our study aimed to investigate CPR performance in COVID-19 ICUs and the effect of ST on the performance of non-ICUp. The hypothesis of the study was largely supported by the findings. The results, based on subjective evaluations by experts (scores on OB1-OB6) and objective mannequin scores (M1 and M3), showed that ICUp performed significantly better than non-ICUp. The observed lack of statistical significance pertains specifically to "chest compression depth" mannequin scores (M2), and this was while controlling for sex. Previous studies indicated that non-ICUp could acquire relevant knowledge to work in ICUs through a 1-day training program (15,16). However, these studies primarily focused on theoretical skills, and there was a lack of clinical performance measures for non-ICUp (15). Our study aimed to fill this gap by including clinical performance measures. Our findings demonstrated that non-ICUp had lower clinical performance in CPR compared with ICUp, highlighting the need for more effective and hands-on clinical training for non-ICUp before their assignment to ICUs.

The hypotheses regarding the lower performance of non-ICUp under ST compared with the no-threat condition and the similar performances of ICUp under ST and no-threat conditions were not supported. Although non-ICUp performed better in the ST condition in some variables, ICUp performed better in "correct place of chest compression" expert scores (OB1) in the no-threat condition, suggesting an ST effect for ICUp rather than non-ICUp. Although the effect was not significant, there was a similar tendency in "chest compression number per minute" mannequin scores (M1), indicating a parallel trend between expert evaluations and mannequin scores to some extent. However, the mannequin may have been more sensitive in evaluating the correct place of chest compression, and it is possible that the experts could not see the exact placement from the camera's perspective.

Contrary to our expectations, the findings also showed that non-ICUp performed better in "effective compression" expert scores (OB3) and "chest compression number per minute" mannequin scores (M1) with marginal significance in the ST condition compared with the no-threat condition. These findings may be related to the difficulty level of "effective compression" and "chest compression number per minute." Previous literature on ST suggested that difficult tasks tended to have stronger effect sizes compared with easy tasks (5). Speculatively, for non-ICUp, "effective compression" and "chest compression number per minute" might have been perceived as easy tasks, and being evaluated and compared with ICUp could have increased their motivation and subsequently improved their performance on these tasks. These unexpected findings call for further research. Considering the inconsistent findings in the literature on ST, it is important to investigate moderating factors (17). Future studies should examine these moderating factors for both ICUp and non-ICUp. For example, the lower scores of ICUp in "correct chest compression place" expert scores (OB1) in the ST condition might be associated with their higher motivation to perform well on this task because CPR is a highly important skill in their specialty compared with non-ICUp. Additionally, it is worth considering the salience of collective identity (e.g., physician identity) and measuring and controlling for its influence in future studies. Moreover, the experiment could be conducted before ICUp and non-ICUp start working together in the ICU because working together against COVID-19 may reduce the ST effect by fostering a sense of collective identity.

Although investigating sex differences in CPR performance was not the primary aim of our study, sex was included as a control variable. The findings showed that male physicians performed better on "appropriate depth of chest compression" expert (OB5) and mannequin (M2) scores, whereas female physicians performed better on "correct chest compression place" mannequin scores (M3). Further research is needed to examine the generalizability and robustness of these sex differences, and if confirmed, medical training programs could consider these differences during the training of male and female physicians.

Our study has several limitations. First, it was conducted in a single center with a limited number of participants. Secondly, the time pressure faced by physicians during the pandemic may have had negative effects on their psychological well-being and performance during the experiment. However, this limitation also reflects the strength of the study because it demonstrates how mandatory changes in organizational structures, such as those implemented during the COVID-19 pandemic, can impact healthcare professionals' performance in future disasters.

The study contributes to the existing literature by comparing the performance of non-ICUp and ICUp physicians during the pandemic. It also combines observational scores and mannequin scores to evaluate CPR performance. The use of mannequins is common in standardizing CPR performance measurements (18,19). Our study demonstrated that the ratings from observers and the mannequin scores were generally aligned with each other. Finally, the study evaluates how physicians are affected by ST in COVID-19 ICUs.

The findings of this study have significant implications for medical education. The observed performance differences between ICUp and non-ICUp underscore the importance of practical experience and exposure to specific clinical settings, such as the ICU, in developing expertise in critical tasks such as CPR. This highlights the need for medical education programs to incorporate sufficient clinical rotations and hands-on training opportunities, allowing students to gain practical skills and experience in real-world scenarios. Research has consistently shown the positive impact of practical training on clinical performance and patient outcomes (20). Furthermore, the interaction between physicians' specialties and ST conditions emphasizes the influence of situational factors on performance. Our study underscores the significance of creating realistic training environments that simulate high-pressure situations because these can enhance the ability of healthcare professionals to perform effectively in challenging circumstances (21,22). Integrating these findings into medical curricula and training programs can ultimately contribute to the development of competent and confident healthcare professionals.

ACKNOWLEDGEMENTS

We would like to thank Dr. Elif Kılıç Çakmak for her contributions to the study.

AUTHOR CONTRIBUTIONS

Conception or design of the work: MT, LK
Data collection: BK
Data analysis and interpretation: MT, BK, LK
Drafting the article: AAY, FU
Critical revision of the article: SBA
The author (MT, LK, FU, AAY, BK, SBA) reviewed the results and approved the final version of the manuscript.

REFERENCES

1. Kite ME, Whitley BE. Psychology of prejudice and discrimination (3rd ed.). Routledge; 2016.
2. Steele CM, Aronson J. Stereotype threat and the intellectual test performance of African Americans. J Pers Soc Psychol 1995;69(5):797-811.
3. Stroessner S, Good C. Stereotype threat: An overview. Retrieved on 2011;9(2):2013.
4. Schmader T, Johns M, Forbes C. An integrated process model of stereotype threat effects on performance. Psychol Rev 2008;115(2):336-56.
5. Shewach OR, Sackett PR, Quint S. Stereotype threat effects in settings with features likely versus unlikely in operational test settings: A meta-analysis. J Appl Psychol 2019;104(12):1514-34.
6. Genel Koronavirüs Tablosu. (2022). Available from: https://covid19.saglik.gov.tr/TR-66935/genel-koronavirus-tablosu.html. Available date: 31/12/2022.
7. Board T. [@tim_n_board]. (2020, March 16). #stayhome #covid [Tweet]. Available from: https://twitter.com/tim_n_board/status/1239499551419047936?s=20
8. Baumgaertner E, Karlamangla S. (2020, March 20). Coronavirus outbreak has doctors, nurses bracing for onslaught. Los Angeles Times. Available from: https://www.latimes.com/california/story/2020-03-20/coronavirus-doctors-nurses-fears-ventilator-icu-emergency.
9. Nguyen, H. H. D., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. Journal of applied psychology, 93(6), 1314.
10. Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Communication methods and measures 2007;1(1):77-89.
11. Krippendorff K. Reliability in content analysis: Some common misconceptions and recommendations. Human communication research 2004;30(3):411-33.
12. Nelson TD. Ageism: Prejudice against our feared future self. Journal of Social Issues 2005;61(2):207-21.
13. Fingerhut AW, Abdou CM. The role of healthcare stereotype threat and social identity threat in LGB health disparities. Journal of Social Issues 2017;73(3):493-507.
14. Burgess DJ, Joseph A, van Ryn M, Carnes M. Does stereotype threat affect women in academic medicine? Acad Med 2012;87(4):506-12.
15. Engberg M, Bonde J, Sigurdsson ST, et al. Training non-intensivist doctors to work with COVID-19 patients in intensive care units. Acta Anaesthesiol Scand 2021;65(5):664-73.
16. Monteverde E, Bosque L, Klappenbach R, et al. Nonintensivist Training to Increase the Staff Capacity of Intensive Care Units During COVID-19 Pandemic Surge in Argentina. Disaster Med Public Health Prep 2021;17:e41.
17. Shapiro J, Aronson J, McGlone MS. Stereotype threat. In: Nelson TD, editor. Handbook of prejudice, stereotyping, and discrimination. Psychology Press; 2009, p. 87-105.
18. Katipoglu B, Madziala MA, Evrin T, et al. How should we teach cardiopulmonary resuscitation? Randomized multi-center study. Cardiol J 2021;28(3):439-45.
19. Ontrup G, Vogel M, Wolf OT, Zahn PK, Kluge A, Hagemann V. Does simulation-based training in medical education need additional stressors? An experimental study. Ergonomics 2020;63(1):80-90.
20. Gordon JA, Hayden EM, Ahmed RA, Pawlowski JB, Khoury KN, Oriol NE. Early bedside care during preclinical medical education: Can technology-enhanced patient simulation advance the Flexnerian ideal? Acad Med 2010;85(2):370-7.
21. Cheng A, Donoghue A, Gilfoyle E, Eppich W. Simulation-based crisis resource management training for pediatric critical care medicine: A review for instructors. Pediatr Crit Care Med 2012;13(2):197-203.
22. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. Med Educ 2012;46(7):636-47.

APPENDIX 1: CPR EVALUATION FORM

Please rate participants' CPR performance from 1 to 4 based on the questions below. (OB: observational)

OB1. Are the chest compressions applied to the correct place?

1. Completely wrong place (ex: abdomen)
2. Correct place - less than 50% of the time
3. Correct place - more than 50% of the time
4. Correct place

OB2. Are the hands placed correctly?

1. Hands/elbows completely wrong (eg: hands made into fists, elbows bent)
2. Hand position correct (but wrong angle or hands are not clasped)
3. Too close to the correct position and angle (the heel of the hand is in full contact and does not separate)
4. Completely in the correct position/angle and does not distort

OB3. Is effective compression applied?

1. Effective in less than 25% of the time
2. Effective in 25-50% of the time
3. Effective in 50-75% of the time
4. Effective in more than 75% of the time

OB4. Is appropriate compression number per minute applied?

1. Too fast or too slow
2. In correct range less than 50% of the time
3. In correct range, more than 50% of the time
4. In correct range all the time

OB5. Is the compression depth appropriate?

1. Too deep or too superficial
2. At the appropriate depth less than 50% of the time
3. At the appropriate depth more than 50% of the time
4. At the appropriate depth

OB6. Is the rib cage allowed to retract?

1. Not allowed
2. Allowed more than 50% of the time
3. Allowed more than 50% of the time
4. Allowed

APPENDIX 2: MANNEQUIN SCORES

Scores of cardiopulmonary resuscitation simulator mannequin (Resusci Anne Simulator®, 'Laerdal Medical'). (M: Mannequin)

M1. Chest compression number per minute (n: 100-120)

1. Too fast or too slow
2. In correct range less than 50% of the time
3. In correct range more than 50% of the time
4. In correct range all the time

M2. Chest compression depth (n: 5-6 cm)

1. Too deep or too superficial
2. At the appropriate depth less than 50% of the time
3. At the appropriate depth more than 50% of the time
4. At the appropriate depth

M3. Are the chest compressions applied to the correct place?

1. Completely wrong place (ex: abdomen)
2. Correct place - less than 50% of the time
3. Correct place - more than 50% of the time
4. Completely correct place indicate