Check for updates

RESEARCH ARTICLE
[version 1; peer review: 1 approved with reservations, 1 not approved]

Do mental health professionals react to verbal aggressiveness differently if it is gender-related? An experimental study employing case-vignettes

[version 1; peer review: 1 approved with reservations, 1 not approved]

Francisco José Eiroa-Orosa[1-3], Cecilia Sanchez-Moscona[1,2]

[1]Section of Personality, Evaluation and Psychological Treatment; Department of Clinical Psychology and Psychobiology, Faculty of Psychology, University of Barcelona, Barcelona, Catalonia, 08035, Spain
[2]First-Person Research Group; Veus, Catalan Federation of 1st Person Mental Health Organisations, Barcelona, Catalonia, Spain
[3]Yale Program for Recovery and Community Health, Department of Psychiatry, Yale School of Medicine, Yale University, New Haven, Connecticut, 06513, USA

v1 **First published:** 03 Feb 2025, **5**:36
https://doi.org/10.12688/openreseurope.19331.1
Latest published: 03 Feb 2025, **5**:36
https://doi.org/10.12688/openreseurope.19331.1

Abstract

Background

Gender bias in the comprehension and management of violent behaviour in mental health settings remains a contentious issue.

Aims

This study examines how gender-related verbal aggressiveness influences the clinical judgment and decision-making processes of mental health professionals (MHPs).

Methods

Four vignettes depicting confrontations between MHPs and service users were created, varying by the service user's gender and whether verbal aggression was gender related. Within an experimental design, a total of 110 MHPs were randomly assigned to vignettes featuring a service user of the opposite gender. Participants completed assessments on clinical judgment, symptomatology, and personal beliefs.

Results

Open Peer Review

Approval Status ✗ ?

	1	2
version 1 03 Feb 2025	✗ view	? view

1. **Johannes Fuss**, University of Duisburg-Essen, Essen, Germany
Maria Krasnova, University of Duisburg-Essen, Essen, Germany
Bastian Nau, University of Duisburg-Essen, Essen, Germany

2. **Emmanuel Biracyaza**, Universite de Montreal, Montreal, Canada

Any reports and responses or comments on the article can be found at the end of the article.

Gender-related verbal aggressiveness influenced clinical judgment, with male professionals demonstrating greater sensitivity to these conditions compared to their female counterparts. Female MHPs exhibited more consistent evaluations across experimental conditions. Gender-specific patterns in symptomatology assignment were observed.

Conclusions

These findings highlight the influence of gender role expectations and sociocultural factors on clinical judgments, which may reinforce gender biases in mental health assessments. Awareness of these biases and targeted training could enhance equitable and unbiased care.

Plain Language Summary

This study looked at how mental health professionals (MHPs) make decisions when dealing with situations where a service user shows verbal aggression. The research focused on whether gender influences how MHPs react to aggressive behaviour. We created four different scenarios where an MHP had to deal with an angry service user. The only difference between the scenarios was the gender of the service user (male or female) and whether the aggression was directed toward the MHP's gender or not.

In total, 110 MHPs participated in the study. They were randomly assigned to one of two scenarios where the service user was of the opposite gender. The findings indicated that professionals' decisions were indeed affected by the experimental conditions, with male professionals showing greater variability in their judgments compared to their female counterparts. However, there were no overarching differences in the severity of judgments based solely on the gender of the MHP or the experimental condition.

The study suggests that societal expectations about gender roles can influence how mental health professionals view and treat mental health issues, sometimes making them think that certain disorders belong to either men or women, which affects their judgments. Understanding and addressing these biases can contribute to more equitable and effective mental healthcare.

Keywords

gender bias; diagnosis; mental health professionals; treatment; verbal aggressiveness; vignette



This article is included in the Marie-Sklodowska-Curie Actions (MSCA) gateway.



This article is included in the Horizon 2020 gateway.



This article is included in the Integrating Gender Analysis into Research collection.

Corresponding author: Francisco José Eiroa-Orosa (fjeiroa@gmail.com)

Author roles: Eiroa-Orosa FJ: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Resources, Validation, Writing – Review & Editing; **Sanchez-Moscona C**: Data Curation, Writing – Original Draft Preparation

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 654808.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2025 Eiroa-Orosa FJ and Sanchez-Moscona C. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Eiroa-Orosa FJ and Sanchez-Moscona C. **Do mental health professionals react to verbal aggressiveness differently if it is gender-related? An experimental study employing case-vignettes [version 1; peer review: 1 approved with reservations, 1 not approved]** Open Research Europe 2025, **5**:36 https://doi.org/10.12688/openreseurope.19331.1

First published: 03 Feb 2025, **5**:36 https://doi.org/10.12688/openreseurope.19331.1

Introduction

Gender bias in mental healthcare has emerged as a contentious topic. Despite the guidelines set forth by ethics codes of practice (e.g. American Psychological Association, 2002; World Medical Association, 2006), mental health professionals are unavoidably susceptible to being influenced by prejudices and stereotypes during their interactions with service users. As a consequence, these biases can lead to biased clinical practices.

Previous research has predominantly focused on investigating gender bias in the diagnosis of personality disorders (McCredie & Morey, 2019), However, there has been relatively less examination of prognosis and treatment recommendations. The findings from these studies have been somewhat inconsistent. In general, the evidence has indicated that women are more prone to receiving mental health diagnoses overall, with a higher likelihood of being diagnosed with conditions such as depression, histrionic personality disorder (Garb, 1997), dependent, and borderline personality disorders (Bjorklund, 2006; Brodie, 2003), whereas men are more likely to be diagnosed with antisocial (Brodie, 2003; Samuel & Widiger, 2009), paranoid, schizoid, schizotypal, narcissistic, and compulsive personality disorders (Brodie, 2003).

Setting diagnoses aside, research indicates that mental health professionals tend to perceive men as mentally healthier but less likely to have favourable prognoses (Garb, 1997). There is also a tendency to perceive men as potentially more violent than women, despite the fact that male and female risk of violence in psychiatric contexts is nearly equal (Nicholls *et al.*, 2004). Moreover, clinicians appear to exhibit higher levels of stigmatising behaviours, such as negative stereotyping or social restrictiveness, when interacting with individuals of the opposite sex (Henderson *et al.*, 2014).

Experimental vignette studies offer a methodological approach to effectively investigate the impact of gender bias on clinicians' decision-making. A vignette consists of a concise narrative depicting a hypothetical situation, which allows researchers to examine respondents' judgments and behaviours in response to the presented scenarios (Atzmüller & Steiner, 2010; Finch, 1987; Sheringham *et al.*, 2021). These vignettes can be presented in various formats, such as text, pictures, videos, or enacted by actors, providing flexibility and realism in the experimental design (Evans *et al.*, 2015). By manipulating key variables while controlling others, vignette studies create a controlled yet lifelike setting, bridging the gap between conventional survey studies and real-life observations (Sheringham *et al.*, 2021).

Case vignettes have proven valuable in experimental psychology for studying various aspects, including the influence of gender bias and aggressiveness on clinical judgment. Studies examining gender bias using vignettes have shown consistent differences in the prevalence rates of personality disorders between men and women. Female subjects are more likely to receive diagnoses of Histrionic Personality Disorder (Erickson, 2001; Samuel & Widiger, 2006) and Borderline Personality Disorder (Delphin, 2001; Giacalone, 1997; Knowles, 2009), while male subjects are more commonly diagnosed with Narcissistic Personality Disorder (Erickson, 2001; Kolker, 1994) and Antisocial Personality Disorder (Samuel & Widiger, 2009).

Findings on depression (Hartman, 2001; Kales *et al.*, 2005; Norman, 2004; Roades, 1994) and sexual disorders (Fuss *et al.*, 2018; Klein *et al.*, 2019) have been mixed. Meta-analytic evidence indicates that men are perceived as more physically and verbally aggressive than women, though the difference narrows when it comes to verbal aggression (Eagly & Steffen, 1986). Accordingly, vignette studies have shown that men are perceived as more dangerous and provoke a greater desire for social distance compared to women (Fuss *et al.*, 2018). Interestingly, when women are depicted as offenders with mental problems, they are considered more deranged and at a higher risk of criminal relapse than men (Sygel *et al.*, 2017). Additionally, men and women exhibit different reactions to various types of aggressive behaviour. Women may respond more strongly to verbal aggressiveness than physical provocation (Bettencourt & Miller, 1996). Accordingly, vignette-based studies reveal that female clinicians perceive clients in a violent context as more dangerous and rate them higher in short-term aggression and violence risk than their male counterparts (McCredie & Morey, 2019; Rojas, 2016).

Despite the existing evidence, the interplay between perceived gender-related aggressiveness and personal values remains unexplored. Addressing this gap, our study aims to assess the impact of mental health service users' gender-related verbal aggressiveness on mental health professionals' decision-making process, specifically during evaluation and diagnosis. We hypothesise that clinicians' clinical judgment will be less favourable towards service users who exhibit gender-related verbal aggressiveness compared to those displaying non-gender-related verbal aggressiveness.

Methods

Design

We conducted an experimental study, which was registered at https://dx.doi.org/10.17605/OSF.IO/25VFK and approved by the University of Barcelona Institutional Review Board [IRB: 00003099] on October 8, 2015. Participants provided informed consent prior to participating in the study. All data were anonymised, and confidentiality was maintained throughout the research process.

The independent variables (IVs) in this study were the service user's gender, which was always opposite to that of the MHP, and the type of verbal aggressiveness displayed, categorised as either gender-related or unrelated to the MHP's gender. These variables were systematically manipulated using clinical vignettes (see extended data at https://dx.doi.org/10.17605/OSF.IO/25VFK). Participants were randomly assigned to one of two vignettes, each featuring a service user of the opposite gender. Consequently, male participants evaluated a vignette involving a female service user, while female participants evaluated one involving a male service user. Both scenarios included instances of verbal aggression, either linked to gender or not.

The Dependent Variables (DVs) included the participants' scores on the Clinical Judgment Scale (CJS; Lee, 2012; Wallach & Strupp, 1960) and their choice of the main vignette symptomatology in a nominal four-level variable (depression, psychosis, obsessions, and mania). Additionally, we considered the following covariates related to the participants: gender, age, profession, years of experience, values (Inglehart, 2008), and scores on the Beliefs and Attitudes towards Service Users' Rights scale (BAMHS; Eiroa-Orosa & Limiñana-Bravo, 2019).

Participants

A minimum of 44 participants per group was considered necessary to detect a statistically significant difference of 0.6 standard deviations or higher. Ultimately, 155 Mental Health Professionals (MHPs) were recruited to participate in the experiment, which took place during the enrolment process of a continuing education course focused on the mental health recovery model (Eiroa-Orosa, 2023). Further recruitment became infeasible due to COVID-19 restrictions on continuing education activities.

Participants who did not complete the entire survey (29%) were excluded from the analysis, resulting in a final sample of 110 participants (59 in the control condition and 51 in the experimental condition). The average age of the participants was 44.71 years, with 87 (79.1%) identifying as women. The majority of the participants were clinical psychologists or clinical psychology residents (n=33; 30.9%), followed by 23 (20.9%) psychiatrists or psychiatry residents, 12 (10.9%) social workers, 11 (10%) social educators, 3 (2.7%) nurses, and 1 (0.9%) occupational therapist. Additionally, 26 participants (23.6%) had other professions, mostly related to direct support. The average years of professional experience among the participants were 14.49. For further details regarding demographic characteristics, please refer to Table 1.

Instruments

Before the course, and upon obtaining written informed consent, participants were requested to complete an online questionnaire. The questionnaire consisted of the Beliefs and Attitudes towards Service Users' Rights scale (BAMHS), followed by sociodemographic information. Subsequently, participants were presented with the randomly assigned clinical vignette and asked a question regarding the main symptomatology of the vignette (mania, depression, obsessions, or psychosis). The Clinical Judgment Scale (CJS) and two values items adapted from the World Values Survey (Inglehart, 2008) were also included in the questionnaire.

Sociodemographic questionnaire

Participants were requested to complete a concise demographic questionnaire, providing information on their birth year, gender, educational level, profession, years of clinical experience, and any personal or relative's history of mental health problems.

Beliefs and Attitudes towards Mental Health Service Users' Rights Scale

The BAMHS (Eiroa-Orosa & Limiñana-Bravo, 2019) is a tool designed to assess mental health professionals' attitudes towards service users'. The scale consists of 25 items, each rated on a 4-point Likert scale. It comprises four dimensions: system criticism / justifying beliefs (α = .701), freedom / coercion (α = .596), empowerment / paternalism (α = .702), and tolerance / discrimination (α = .620). The system criticism / justifying beliefs dimension captures mental health professionals' beliefs that justify the status quo of service users. The second dimension encompasses recurring topics in discussions about service users' rights. The empowerment / paternalism subscale reflects beliefs concerning the perceived inability of individuals diagnosed with a mental disorder to take charge of their own lives. The final dimension collects prevailing prejudices towards mental health service users. In our study, higher scores indicate a higher level of rights violations.

Clinical Judgment Scale

The CJS (Lee, 2012; Wallach & Strupp, 1960) is an 18-item questionnaire employing a 7-point Likert scale for each item. It assesses therapists' overall clinical impressions of a clinical case on a pessimism-optimism dimension concerning clinical prognosis. Higher scores on the CJS indicate a poorer prognosis. Throughout the questionnaire, participants are prompted to make judgments on various factors: Poor Skills (α = .711), Severity (α = .716), Motivation (α = .756) and Treatment (α = .807). The Poor Skills dimension evaluates the perception of service users' social and relational skills, whereas Severity assesses how therapists perceive the service users' insight into their problems, their contact with reality, and their prognosis without treatment. Motivation reflects the professionals' level of interest in working with the service user, while Treatment measures the professional's impression of the service users' readiness for treatment and its potential effectiveness.

Values items adapted from the World Values Survey

Two items from the World Values Survey (Inglehart, 2008) were included in the study to assess participants' alignment with either traditional or progressive values. These items were rated on a 6-point Likert scale, ranging from "I completely disagree" to "I completely agree". The items were as follows: "Currently there is total equality between men and women" and "A child needs a mother to grow up happy".

Procedure

The instruments and vignettes were administered through the Qualtrics survey platform. Four clinical vignettes (available as extended data) were specifically designed for this study. All four vignettes presented identical symptomatology, but each one varied in terms of the service user's gender (male or female) and the type of verbal aggressiveness displayed (related to the opposite gender or unrelated to gender entirely).

In each vignette, the scenario depicted a conversation with a service user who appeared disoriented upon admission to an acute crisis unit. During the interaction, the service user exhibited verbally aggressive behaviour towards the mental health professional. In the vignettes where verbal aggressiveness was not related to gender, the service user's family described the individual as reserved. However, in the vignette with a male

Table 1. Sociodemographic characteristics and baseline scores on the BAMHS scale.

	Female participants							Male participants							Statistical Significance					
	Intervention (n=41)		Control (n=46)		Statistical Significance			Intervention (n=10)		Control (n=13)		Statistical Significance			Intervention vs. Control			Women vs. Men		
	N	%	N	%	OR, *95%CI*		*p*	N	%	N	%	OR,*95%CI*		*p*	OR, 95%CI		*p*	OR. 95%CI		*p*
Higher education*	30	73.2	38	82.6	1.742, 0.622-4.873		.288	9	90	9	69.2	.250, 0.023-2.696		.231	1.205, 0.487-2.981		0.686	0.994, 0.326-3.028		0.992
	M	SD	M	SD	t	*p*	d	M	SD	M	SD	z	*p*	d	t	*p*	d	t	*p*	d
Age	44.00	10.29	44.13	10.82	-0.057	.954	-0.01	45.50	10.88	48.38	9.55	-0.435	.663	-0.28	-0.386	.700	-0.07	-1.254	.213	-0.29
Years of experience as MHP	13.33	8.77	13.86	9.51	-0.239	.812	-0.06	17.75	7.89	18.00	9.31	-0.045	.964	-0.03	-0.288	.774	-0.06	-1.802	.075	-0.48
VALUES																				
Equality between men and women	2.41	1.00	2.22	0.87	0.986	.327	0.21	2.30	0.95	2.54	1.27	-0.295	.768	-0.21	0.559	.577	0.11	-0.545	.587	-0.13
A kid needs a mother to be happy	3.12	1.23	3.50	1.11	-1.507	.135	-0.32	3.70	1.49	3.46	1.33	-0.543	.587	0.17	-1.101	.273	-0.21	-0.851	.397	-0.20
BAMHS																				
Beliefs	2.47	0.44	2.35	0.37	1.344	.183	0.29	2.43	0.26	2.48	0.34	-0.5	.617	-0.18	1.083	.281	0.21	-0.582	.561	-0.14
Coercion	2.50	0.43	2.33	0.39	1.9	.061	0.41	2.56	0.35	2.35	0.41	-1.327	.185	0.54	2.272	.025	0.44	-0.385	.701	-0.09
Paternalism	2.43	0.47	2.21	0.37	2.366	.020*	0.51	2.10	0.49	2.51	0.49	-1.86	.063	-0.83	0.973	.333	0.19	-0.162	.872	-0.04
Discrimination	1.90	0.45	1.84	0.40	0.65	.517	0.14	1.68	0.41	1.79	0.43	-0.447	.655	-0.27	0.32	.749	0.06	1.324	.188	0.31
Total	2.38	0.33	2.25	0.28	1.94	.056	0.42	2.27	0.14	2.39	0.29	-1.338	.181	-0.53	1.322	.189	0.25	-0.364	.716	-0.09
CJS																				
Treatment	3.84	0.59	3.90	0.73	-0.451	.653	-0.10	3.40	0.45	4.09	0.67	-2.746	.006**	-1.18	-1.521	.131	-0.29	0.524	.601	0.12
Motivation	2.80	0.78	2.79	0.71	0.043	.966	0.01	2.93	1.07	2.97	0.62	-0.535	.592	-0.05	-0.048	.962	-0.01	-0.921	.359	-0.22
Severity	5.07	0.73	4.66	0.81	2.489	.015*	0.54	4.22	1.05	4.88	0.72	-1.527	.127	-0.75	1.254	.213	0.24	1.348	.181	0.32
Poor skills	4.32	0.47	4.27	0.52	0.533	.595	0.12	3.65	0.75	4.40	0.48	-2.476	.013*	-1.23	-1.003	.318	-0.19	1.698	.092	0.40
CJS total	4.11	0.41	4.01	0.55	1.027	.308	0.22	3.61	0.42	4.19	0.37	-2.981	.003**	-1.49	-0.359	.72	-0.07	1.046	.298	0.25

Note. Percentage of participants with at least a basic university degree. *p <.05, **p <0.01, ***p <.001. BAMHS: Beliefs and Attitudes towards Service Users' Rights scale CJS: Clinical Judgment Scale.

Open Research Europe 2025, 5:36 Last updated: 28 MAY 2025

service user displaying gender-related aggressiveness, the family portrayed him as a very traditional man. Conversely, in the vignette with a female service user exhibiting gender-related aggressiveness, the family described her as a very modern woman (for a comprehensive vignette description, refer to the extended data).

The creation of the vignettes adhered to Heverly *et al.* (1984) four-step protocol to enhance its validity: 1) to identify the experimental factors to be varied, 2) to generate descriptions of behaviour that clearly reflect concrete levels of the desired factor, 3) to empirically validate these descriptions, and 4) to construct the whole vignette from the validated pieces.

Statistical analyses

We employed percentages, means, and standard deviations to describe socio-demographic information, the values items adapted from Inglehart (2008), the BAMHS and the CJS scores. Due to the small size of the male sample, non-parametric tests were used for tests within this group. To explore differences between the experimental and control groups, as well as between male and female participants, t-tests or Mann-Whitney U tests and Cohen's d were applied for quantitative variables. For categorical variables, chi-square and odds-ratio tests were used. Analyses involved multiple group comparisons using different combinations of the levels of the independent variable, such as gender-neutral men attacking a female MHP vs. gender-related men attacking a female MHP, gender-neutral attacking a male MHP vs. gender-related women attacking a male MHP, men vs. women, and verbal aggressiveness related to gender vs. non-gender-neutral verbal aggressiveness. Additionally, Pearson correlations were performed among the values, BAMHS, and CJS scores. For statistically significant correlations, ANCOVAs were conducted to test the main effects and interaction of the experimental condition on the CJS scores while controlling for covariates.

Results

Sociodemographic characteristics, values, and BAMHS scores are presented in Table 1. Statistically significant differences were observed in Paternalism scores between women in the experimental and control groups, as well as in Coercion scores between the experimental and control groups.

CJS scores are also provided in Table 1. A statistically significant difference was detected in the Severity score between women in the control and intervention groups. Regarding men, statistically significant differences were found in the Treatment and Poor Skills subscales, as well as in the total CJS score, between the intervention and control groups.

Data on the main symptomatology assigned to the vignettes can be found in Table 2. Our findings reveal statistically significant differences between male and female participants concerning psychosis. Female MHPs tended to assign psychosis as the main symptomatology (67.8%), irrespective of whether they belonged to the experimental or control group. In contrast, most male MHPs designated obsessions as the main symptomatology of the vignettes (56.5%), with no statistically significant differences observed between the experimental groups.

When conducting Pearson correlations between CJS, BAMHS, and values scores, statistically significant relationships were observed between the CJS Total score and the BAMHS Total and Coercion scores. Furthermore, statistically significant correlations were found between CJS Treatment and BAMHS Coercion scores, as well as between CJS Severity and BAMHS Total, Beliefs, Coercion, and Paternalism scores (see Table 3). Subsequently, the results of the ANCOVAs with the experimental condition as a factor and BAMHS subscales with statistically significant correlations with CJS scores are presented in Table 4. In all cases, the covariate was found to be statistically significant. Notably, in the model including

Table 2. Symptomatology assigned to the vignette.

	Female participants					Male participants					Women vs. Men	Intervention vs. Control
	Intervention (n=47)		Control (n=46)		OR. 95%IC	Intervention (n=10)		Control (n=16)		OR. 95%IC	OR. 95%IC	OR. 95%IC
	N	%	N	%		N	%	N	%			
Depression	2	4.9	3	6.5	NC	3	30	1	7.7	NC	3.453, 0.846-14.090	1.495, 0.379-5.893
Psychosis	29	70.7	30	65.2	1.289. 0.521-3.189	3	30	3	23.1	1.429. 0.220-9.262	0.167, 0.60-0.471***	1.327, 0.617-2.854
Obsessions	9	22.0	12	26.1	0.797. 0.269-2.144	4	40	9	69.2	0.296, 0.053-1.669	4.086, 1.565-10.666**	0.619, 0.271-1.413
Mania	1	2.4	1	2.2	NC	0	0	0	0	NC	NC	NC

Note. NC= Not Calculated due to boxes with n<3. *p <.05, **p <0.01, ***p <.001.

Table 3. Pearson's correlation between the BAMHS and the CJS.

	Beliefs	Coercion	Paternalism	Discrimination	BAMHS Total	Equality between men and women	A kid needs a mother to be happy
Treatment	-0.064	0.292**	0.042	0.167	0.099	-0.011	0.051
Motivation	0,008	0,11	0,09	0,114	0,14	-0,024	0,124
Severity	,315**	,211*	,219*	0,119	,293**	0,064	-0,021
Poor skills	0,1	0,062	0,046	0,035	0,069	-0,003	-0,149
CJS Total	0,144	,273**	0,156	0,168	,234*	0,018	0,008

Note. *p<.05. **p<.01. ***p<.001. BAMHS: Beliefs and Attitudes towards Service Users' Rights scale CJS: Clinical Judgment Scale.

Table 4. ANCOVA using CJS as dependent variable, experimental condition as factor and BAMHS scores as covariate.

Dependent Variable	Covariate	F	Sig.	Partial Eta Squared	Adjusted R Squared
Treatment	Coercion	13.255	<.0001	0.111	.112
	Condition	5.264	.024	0.047	
Severity	Beliefs	10.985	<.001	0.094	.089
	Condition	0.779	.379	0.007	
	Coercion	4.061	.046	0.037	.031
	Condition	0.543	.463	0.005	
	Paternalism	4.958	.028	0.044	.040
	Condition	1.130	.290	0.010	
	BAMHS Total	9.239	.003	0.079	.076
	Condition	0.822	.367	0.008	
CJS Total	BAMHS Total	6.572	.012	0.058	.041
	Condition	0.465	.492	0.004	

Note. BAMHS: Beliefs and Attitudes towards Service Users' Rights scale CJS: Clinical Judgment Scale.

CJS Treatment and BAMHS Coercion, the experimental condition displayed statistical significance, with this model also exhibiting a higher adjusted R-squared. Additionally, when adding the interaction between the covariate (BAMHS) and the experimental group, statistically significant results were only obtained for the BAMHS Paternalism with the CJS Severity ($F = 4.485$, $p = .037$, $\eta_p^2 = 0.041$).

Discussion

In this experimental study, we explored the potential differential impact of gender-related verbal aggressiveness on mental health professionals' assessment of a clinical vignette. Our findings revealed that such influence exists to a certain extent, and there is a correlation between participants' attitudes and beliefs regarding coercive practices and how the experiment affected their clinical judgments concerning treatment.

The results partially support our main hypothesis that professionals' judgment would be less favourable for vignettes depicting service users whose verbal aggressiveness was related to gender compared to those whose aggressiveness was not. However, it is important to note that these differences can only be observed within each gender and not globally across all participants.

Aggressiveness may be viewed as a direct consequence of gender roles. Previous experimental studies have consistently shown that men tend to be perceived as more aggressive than women (e.g., Eagly & Steffen, 1986), unless gender role features are less salient in the given circumstances. Reciprocity plays a significant role in social interactions, making provocation a common justification for aggression. For women, provocation becomes particularly relevant as it allows them to break free from the limitations often imposed by gender role norms (Bettencourt & Miller, 1996). In addition, differences in aggression between provoked and neutral conditions are more pronounced for women than for men (Bettencourt & Miller, 1996; McCredie & Morey, 2019; Rojas, 2016). Numerous studies, incorporating variations in participants' gender and levels of provocation, consistently demonstrate that women exhibit less aggression than men under relatively neutral conditions. However, when provoked, women tend to display slightly higher levels of aggression compared to men (Schuck *et al.*, 1971; Taylor & Epstein, 1967). In our study, we found that male participants exhibited more significant differences between the experimental conditions compared to female participants.

The observation that female participants in our study exhibited more stable clinical judgments than male participants can be attributed to various factors. Firstly, previous studies have consistently found that males tend to engage in riskier behaviours compared to females (e.g. Pawlowski *et al.*, 2008). Secondly, women are inclined to downplay hostility and conflict (Prakash & Caeli Flores, 1985) and tend to score higher in agreeableness, which encompasses traits like empathy, kindness, and the maintenance of social harmony, in comparison to men (Weisberg *et al.*, 2011). These higher rates of agreeableness in women may lead them to respond in ways that align with socially accepted norms, while men may feel more inclined to take risks when making decisions. In our study, this could translate to female participants making less severe clinical judgments compared to male participants.

Research on gender differences in stability when making judgments has yielded mixed results. One fMRI study indicated differences in Emotional Stability between men and women, suggesting that men exhibit moderately more Emotional Stability than women across the countries included in the sample (McRae *et al.*, 2008). Accordingly, some studies have reported that women tend to display higher neuroticism, which is associated with less emotional stability (e.g. Grijalva *et al.*, 2015). However, in contrast to these findings, another study found no stable gender differences in individual confidence on judgments, with the stability depending on the knowledge domain or cognitive processes activated (Jonsson & Allwood, 2003). In contrast, in our study, we observed that female participants showed more stability when making clinical judgments compared to male participants.

Becker and Wagner (2009) propose that different types of gender identity can be partly predicted by varying levels of women's endorsement of sexist beliefs. Women who strongly identify with the 'female' category tend to internalise more progressive identity contents and reject benevolent, hostile, and modern sexism. Additionally, their study highlights the potential interaction between gender roles and gender identity. Female participants' responses to a text describing a woman whose gender role varied (from feminist to traditional) depended on their individual level of gender identification (Becker & Wagner, 2009). However, in our study, no statistically significant correlations were found between female participants with progressive or traditional values and the clinical judgments made for both experimental conditions.

The prevalence of specific diagnoses varied by gender in our study. These differences could be linked to the actual prevalence of these disorders, which may influence professionals' stereotypes. Accordingly, several studies have indicated that the incidence of psychosis is higher in men compared to women. Moreover, findings on symptoms are inconclusive, but it appears that men experience more negative symptoms, while women tend to have more affective symptoms (Ochoa *et al.*, 2012). The incidence of psychosis is greater for men when negative symptoms and a longer duration of the illness (associated with worse prognosis) are described, whereas there is an equivalent risk for both sexes when affective symptoms and a shorter duration of the illness (related to a better prognosis) are present. Consistently, our results demonstrate that more men were assigned to the psychosis symptomatology category than women. Similarly, although obsessive disorders may be more common among males during childhood, they become more prevalent among females during adolescence and adulthood (Mathes *et al.*, 2019).

As expected in a vignette study, our results revealed that clinicians' pre-existing beliefs had a stronger predictive power on their clinical judgments compared to the experimental condition. Specifically, our findings indicate that the BAMHS score may predict the severity of participants' assessment of the given vignette. This suggests that the BAMHS not only serves as a valuable tool to measure the impact of recovery and anti-stigma interventions carried out with mental health professionals (Eiroa-Orosa, 2023; Eiroa-Orosa *et al.*, 2021), but also holds the potential to predict the type of decisions a professional may make.

Furthermore, these results highlight the significance of clinicians' attitudes and beliefs when evaluating a service user. Prior research has demonstrated a tendency among healthcare providers to hold pessimistic views regarding service users' reality and their chances of recovery (Henderson *et al.*, 2014; Schulze, 2007). Such pessimism can contribute to stigma and act as a barrier to recovery (Knaak *et al.*, 2017).

Similarly, the use of coercive measures remains prevalent in mental health care. While a significant number of MHPs view coercion as a violation of fundamental rights, a considerable percentage consider it an indispensable measure in psychiatry and beneficial for service users (Morandi *et al.*, 2021). Furthermore, the more a professional is involved in situations where coercion is employed, the more they tend to believe in its

efficacy (van Doeselaar *et al.*, 2008) and perceive it as a means of achieving better care and providing increased security (Dahan *et al.*, 2018). Consequently, the results of this study suggest that professionals who endorse coercive practices tend to make more pessimistic judgments.

This study has certain limitations that should be considered. Firstly, the small size of the male sample has resulted in insufficient statistical power for gender-disaggregated analyses. Additionally, data on ethnic and cultural identification could not be collected due to the low ethnic-cultural diversity of mental health professionals, which could have compromised participant anonymity. Regarding the vignettes, they do not fully meet all the factors in the quality assessment framework designed by Sheringham *et al.* (2021) for evaluating the robustness of experimental vignette studies. Only one vignette was presented to each participant, which prevented the assessment of intra-subject judgments' variation (number). Moreover, no validity or credibility assessment was conducted to confirm the vignettes' similarity to real-life scenarios (assessment). However, the study did have some strengths in the design of the vignettes. A comprehensive description of the vignette construction was provided (credibility), the vignettes included variability of the experimental factors (variability), the textual modality was followed (modality), and a full description of the vignettes used in the study was available (description). Regarding the study design, although the purpose of the study was concealed from participants, no active strategies were implemented to further hide the purpose, such as including a distracting element in the vignette (concealment). Additionally, the study did not control for whether the completion of questionnaires took place in contexts that replicated aspects of healthcare (realism). However, the sampling and response rate were clearly explained (sampling and response rate). Finally, as only one vignette was provided per participant, the vignette was not included as a covariate to control for in the analysis (analysis).

It is important to note that the experiment was designed differently for male and female participants. In the experimental vignettes, we depicted a "traditional" male service user and a "modern" female service user, leading to the use of different verbal aggressiveness. However, this approach was inevitable as the vignettes were specifically designed to analyse participants' reactions after being provoked in relation to their gender. Given the existence of real social differences, gender-related verbal aggressiveness holds distinct connotations for men and women. Moreover, considering that women generally hold more pro-feminist ideas than men (Rojas, 2016), we can expect that reactions to a progressive person would differ from those to a traditional person for both genders. Similarly, in both control conditions, a reserved person was described, but the interpretation of this trait might differ between genders. Therefore, it is essential to acknowledge that the intervention and control conditions are not directly comparable between genders as they are not entirely equivalent. These differences were necessary to examine the unique aspects of verbal aggressiveness related to gender and its impact on the participants' judgments.

Despite its limitations, this study sheds light on a highly contentious issue. The findings underscore the need for significant changes in the approach of MHPs to provide unbiased care in line with ethical codes. Recognising the potential impact of gender role expectations can help creating a safer therapeutic environment. To address this, establishing open discussions, offering continuing education, and providing supervision on the influence of gender roles in treatment can be valuable tools. Additionally, adopting dimensional diagnostic approaches may prove beneficial. Further research is essential to expand our understanding of how gender influences decision-making in mental health care and to effectively mitigate gender bias.

Ethics and consent

We conducted an experimental study, which was registered at https://dx.doi.org/10.17605/OSF.IO/25VFK and approved by the University of Barcelona Institutional Review Board [IRB: 00003099] on October 8, 2015. Participants provided informed consent prior to participating in the study. All data were anonymised, and confidentiality was maintained throughout the research process.

Before the course, and upon obtaining written informed consent, participants were requested to complete an online questionnaire.

Data availability

Underlying data

Open Science Foundation - Gender bias and personal values influence on mental health professionals https://dx.doi.org/10.17605/OSF.IO/25VFK (Eiroa-Orosa & Sanchez-Moscona, 2025).

This project contains the following underlying data:

- Database.
- Syntax.

Data are available under the terms of the CC-By Attribution 4.0 International

Extended data

Open Science Foundation - Gender bias and personal values influence on mental health professionals https://dx.doi.org/10.17605/OSF.IO/25VFK (Eiroa-Orosa & Sanchez-Moscona, 2025).

This project contains the following extended data:

- Vignettes used in the experiment.

Data are available under the terms of the CC-By Attribution 4.0 International

References

American Psychological Association: **Ethical principles of psychologists and code of conduct.** *Am Psychol.* 2002; **57**(12): 1060–1073.
PubMed Abstract | Publisher Full Text

Atzmüller C, Steiner PM: **Experimental vignette studies in survey research.** *Methodology.* 2010; **6**(3): 128–138.
Publisher Full Text

Becker JC, Wagner U: **Doing gender differently-The interplay of strength of gender identification and content of gender identity in predicting women's endorsement of sexist beliefs.** *Eur J Soc Psychol.* 2009; **39**(4): 487–508.
Publisher Full Text

Bettencourt BA, Miller N: **Gender differences in aggression as a function of provocation: a meta-analysis.** *Psychol Bull.* 1996; **119**(3): 422–447.
PubMed Abstract | Publisher Full Text

Bjorklund P: **No man's land: gender bias and social constructivism in the diagnosis of borderline personality disorder.** *Issues Ment Health Nurs.* 2006; **27**(1): 3–23.
PubMed Abstract | Publisher Full Text

Brodie RE: **Race, sex, and class bias in the diagnosis of DSM -IV disorders.** University of California, Santa Barbara, 2003.

Dahan S, Levi G, Behrbalk P, *et al.*: **The impact of 'being there': psychiatric staff attitudes on the use of restraint.** *Psychiatr Q.* 2018; **89**(1): 191–199.
PubMed Abstract | Publisher Full Text

Delphin ME: **Gender and ethnic bias in the diagnosis of Antisocial and borderline personality disorders.** Purdue University, 2001.
Reference Source

Eagly AH, Steffen VJ: **Gender and aggressive behavior: a meta-analytic review of the social psychological literature.** *Psychol Bull.* 1986; **100**(3): 309–330.
PubMed Abstract | Publisher Full Text

Eiroa-Orosa FJ: **Beyond recovery: toward rights-based mental health care — A cluster randomized wait-list controlled trial of a recovery and rights training for mental health professionals with or without first person accounts.** *Front Psychol.* 2023; **14**: 1152581.
PubMed Abstract | Publisher Full Text | Free Full Text

Eiroa-Orosa FJ, Limiñana-Bravo L: **An instrument to measure mental health professionals' beliefs and attitudes towards service users' rights.** *Int J Environ Res Public Health.* 2019; **16**(2): 244.
PubMed Abstract | Publisher Full Text | Free Full Text

Eiroa-Orosa FJ, Lomascolo M, Tosas-Fernández A: **Efficacy of an intervention to reduce stigma beliefs and attitudes among primary care and mental health professionals: two cluster randomised-controlled trials.** *Int J Environ Res Public Health.* 2021; **18**(3): 1214.
PubMed Abstract | Publisher Full Text | Free Full Text

Eiroa-Orosa FJ, Sanchez-Moscona C: **Gender bias and personal values influence on mental health professionals.** OSF, 13 Jan, 2025.
http://www.doi.org/10.17605/OSF.IO/25VFK

Erickson KB: **Psychologist gender and sex bias in diagnosing histrionic and narcissistic personality disorders.** Indiana State University, 2001.
Reference Source

Evans SC, Roberts MC, Keeley JW, *et al.*: **Vignette methodologies for studying clinicians' decision-making: validity, utility, and application in ICD-11 field studies.** *Int J Clin Health Psychol.* 2015; **15**(2): 160–170.
PubMed Abstract | Publisher Full Text | Free Full Text

Finch J: **The vignette technique in survey research.** *Sociology.* 1987; **21**(1): 105–114.
Publisher Full Text

Fuss J, Briken P, Klein V: **Gender bias in clinicians' pathologization of atypical sexuality: a Randomized Controlled Trial with mental health professionals.** *Sci Rep.* 2018; **8**(1): 3715.
PubMed Abstract | Publisher Full Text | Free Full Text

Garb HN: **Race bias, social class bias, and gender bias in clinical judgment.** *Clinical Psychology: Science and Practice.* 1997; **4**(2): 99–120.
Publisher Full Text

Giacalone RC: **A study of clinicians' attitudes and sex bias in the diagnosis of borderline personality disorder and posttraumatic stress disorder.** The Wright Institute, 1997.
Reference Source

Grijalva E, Newman DA, Tay L, *et al.*: **Gender differences in narcissism: a meta-analytic review.** *Psychol Bull.* 2015; **141**(2): 261–310.
PubMed Abstract | Publisher Full Text

Hartman JS: **Are clinicians biased? The role of client variables in clinician assessment and diagnosis of depressive disorders.** University of Cincinnati, 2001. .
Reference Source

Henderson C, Noblett J, Parke H, *et al.*: **Mental health-related stigma in health care and mental health-care settings.** *Lancet Psychiatry.* 2014; **1**(6): 467–482.
PubMed Abstract | Publisher Full Text

Heverly MA, Fitt DX, Newman FL: **Constructing case vignettes for evaluating clinical judgment: an empirical model.** *Evaluation and Program Planning.* 1984; **7**(1): 45–55.
Publisher Full Text

Inglehart RF: **Changing values among Western Publics from 1970 to 2006.** *West European Politics.* 2008; **31**(1–2): 130–146.
Publisher Full Text

Jonsson AC, Allwood CM: **Stability and variability in the realism of confidence judgments over time, content domain, and gender.** *Personality and Individual Differences.* 2003; **34**(4): 559–574.
Publisher Full Text

Kales HC, Neighbors HW, Blow FC, *et al.*: **Race, gender, and psychiatrists' diagnosis and treatment of major depression among elderly patients.** *Psychiatr Serv.* 2005; **56**(6): 721–728.
PubMed Abstract | Publisher Full Text

Klein V, Briken P, Schröder J, *et al.*: **Mental health professionals' pathologization of compulsive sexual behavior: does clients' gender and sexual orientation matter?** *J Abnorm Psychol.* 2019; **128**(5): 465–472.
PubMed Abstract | Publisher Full Text

Knaak S, Mantler E, Szeto A: **Mental illness-related stigma in healthcare: barriers to access and care and evidence-based solutions.** *Healthc Manage Forum.* 2017; **30**(2): 111–116.
PubMed Abstract | Publisher Full Text | Free Full Text

Knowles A: **Disorders of extreme stress, not otherwise specified, posttraumatic stress disorder, and borderline personality disorder: a vignette study exploring clinicians' diagnostic perceptions.** Indiana State University, 2009.
Reference Source

Kolker JI: **An investigation of diagnostic sex bias for narcissistic personality disorder, in comparison to histrionic and antisocial personality disorders.** The University of North Carolina at Greensboro, 1994.
Reference Source

Lee Y: **Exploring the impact of Asian stereotype endorsement, multicultural counseling competence, and motivation to respond without prejudice on White therapists' clinical judgment.** [Columbia University], 2012.
Publisher Full Text

Mathes BM, Morabito DM, Schmidt NB: **Epidemiological and clinical gender differences in OCD.** *Curr Psychiatry Rep.* 2019; **21**(5): 36.
PubMed Abstract | Publisher Full Text

McCredie MN, Morey LC: **An investigation of the influence of case and clinician gender in clinical decision-making.** *Personal Disord.* 2019; **10**(3): 286–290.
PubMed Abstract | Publisher Full Text

McRae K, Ochsner KN, Mauss IB, *et al.*: **Gender differences in emotion regulation: an fMRI study of cognitive reappraisal.** *Group Process Intergroup Relat.* 2008; **11**(2): 143–162.
PubMed Abstract | Publisher Full Text | Free Full Text

Morandi S, Silva B, Mendez Rubio M, *et al.*: **Mental health professionals' feelings and attitudes towards coercion.** *Int J Law Psychiatry.* 2021; **74**: 101665.
PubMed Abstract | Publisher Full Text

Nicholls TL, Ogloff JRP, Douglas KS: **Assessing risk for violence among male and female civil psychiatric patients: the HCR-20, PCL:SV, and VSC.** *Behav Sci Law.* 2004; **22**(1): 127–158.
PubMed Abstract | Publisher Full Text

Norman J: **Gender bias in the diagnosis and treatment of depression.** *Int J Ment Health.* 2004; **33**(2): 32–43.
Reference Source

Ochoa S, Usall J, Cobo J, *et al.*: **Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review.** *Schizophr Res Treatment.* 2012; **2012**(1): **916198.**
PubMed Abstract | Publisher Full Text | Free Full Text

Pawlowski B, Atwal R, Dunbar RIM: **Sex differences in everyday risk-taking behavior in humans.** *Evolutionary Psychology.* 2008; **6**(1): 147470490800600.
Reference Source

Prakash V, Caeli Flores R: **A study of psychological gender differences: applications for advertising format.** *NA - Advances in Consumer Research.* 1985; **12**: 231–237.
Reference Source

Roades LA: **Gender and race bias in the diagnosis of major depressive episode and alcohol dependence.** University of Missouri, 1994.
Reference Source

Rojas EG: **Gender bias and clinical judgment: examining the influence of attitudes toward women on clinician perceptions of dangerousness.** Columbia University, 2016.

Samuel DB, Widiger TA: **Clinicians' judgments of clinical utility: a comparison of the DSM-IV and five-factor models.** *J Abnorm Psychol.* 2006; **115**(2): 298–308.
PubMed Abstract | **Publisher Full Text**

Samuel DB, Widiger TA: **Comparative gender biases in models of personality disorder.** *Personality and Mental Health.* 2009; **3**(1): 12–25.
Publisher Full Text

Schuck SZ, Schuck A, Hallam E, *et al.*: **Sex differences in aggressive behavior subsequent to listening to a radio broadcast of violence.** *Psychol Rep.* 1971; **28**(3): 931–936.
Publisher Full Text

Schulze B: **Stigma and mental health professionals: a review of the evidence on an intricate relationship.** *Int Rev Psychiatry.* 2007; **19**(2): 137–155.
PubMed Abstract | **Publisher Full Text**

Sheringham J, Kuhn I, Burt J: **The use of experimental vignette studies to identify drivers of variations in the delivery of health care: a scoping review.** *BMC Med Res Methodol.* 2021; **21**(1): 81.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Sygel K, Sturup J, Fors U, *et al.*: **The effect of gender on the outcome of forensic psychiatric assessment in Sweden: a case vignette study.** *Crim Behav Ment Health.* 2017; **27**(2): 124–135.
PubMed Abstract | **Publisher Full Text**

Taylor SP, Epstein S: **Aggression as a function of the interaction of the sex of the aggressor and the sex of the victim.** *J Pers.* 1967; **35**(3): 474–486.
PubMed Abstract | **Publisher Full Text**

van Doeselaar M, Sleegers P, Hutschemaekers G: **Professionals' attitudes toward reducing restraint: the case of Seclusion in The Netherlands.** *Psychiatr Q.* 2008; **79**(2): 97–109.
PubMed Abstract | **Publisher Full Text**

Wallach MS, Strupp HH: **Psychotherapists' clinical judgments and attitudes toward patients.** *J Consult Psychol.* 1960; **24**(4): 316–323.
Publisher Full Text

Weisberg YJ, Deyoung CG, Hirsh JB: **Gender differences in personality across the ten aspects of the big five.** *Front Psychol.* 2011; **2**: 178.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

World Medical Association: **WMA International code of medical ethics**. 2006.
Reference Source

Open Peer Review

Current Peer Review Status: ✗ ?

Version 1

Reviewer Report 28 May 2025

https://doi.org/10.21956/openreseurope.20921.r52977

© **2025 Biracyaza E.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The author(s) is/are employees of the US Government and therefore domestic copyright protection in USA does not apply to this work. The work may be protected under the copyright laws of other jurisdictions when used in those jurisdictions.

?

Emmanuel Biracyaza
Universite de Montreal, Montreal, Québec, Canada

Reviewer report

Is the work clearly and accurately presented, and does it cite the current literature?

I find that the research is interesting since it addresses a timely and highly relevant topic. In today's climate where gender equity and bias in mental health care are central to research and practice, this study explores an important dimension: how gender-related verbal aggressiveness influences mental health professionals' clinical judgment. The strengths include that the abstract effectively summarizes the study's rationale, aims, design, and main findings, providing a clear and focused overview. However, some parts like background, methods and results should be expanded. The introduction offers a solid background on gender bias in clinical diagnosis and treatment perceptions, establishing the broader significance of the study. The use of case vignettes is methodologically appropriate for assessing implicit clinician responses in a controlled but realistic format. Although this study has strengths, it has several areas that require improvement:

1. One of the main weaknesses in the introduction is the heavy reliance on outdated references, with many key studies cited from the 1990s and early 2010s. To ensure the study reflects the current understanding of gender bias, aggression, and mental health practice, the authors should incorporate more recent literature (from the past five years).
2. In some sections, the authors suggest that men experience more issues or are judged more harshly than women (e.g., greater perceived dangerousness, worse prognoses). However, the factors contributing to these differences are not adequately explained. The authors should explore why these disparities might exist—whether due to cultural stereotypes, clinician bias, or systemic influences. This would enhance the depth of the analysis and improve the theoretical coherence of the introduction.
3. The introduction could be improved by engaging more deeply with recent developments in gender discourse, including changes in mental health policy, increasing awareness of gender fluidity, or contemporary research on implicit bias. This would make the argument more relevant and grounded in current practice.

4. While the general research context is well presented, the specific research gap addressed namely, the impact of gender-related verbal aggressiveness on clinical decision-making—is not clearly and explicitly stated. A sharper articulation of the research question would strengthen the logical flow of the background.
5. Although the hypothesis is implied in the abstract and introduction, it would benefit from being clearly and directly stated. Explicitly presenting the hypothesis would help readers better understand the study's focus and expectations.

Is the study design appropriate and does the work have academic merit?

There are several areas that need revision or clarifications:

1. The authors mention a minimum of 44 participants per group, but it's not clear whether this calculation was based on a power analysis using expected effect sizes and significance levels. I suggest the authors providing more detail on how the power calculation was conducted, especially given the complexity of multiple comparisons and interactions. In addition to this, the males were underrepresented which affects the results and conclusions.
2. Although vignettes were designed to manipulate gender-related vs. neutral verbal aggressiveness, the specific content and validation results of these manipulations are not provided in the main text (only referenced via extended data). I recommend the authors include brief examples or summaries of the vignettes and validation process in the main text to demonstrate the construct validity of the manipulations.
3. The study defines the independent variables as verbal aggressiveness (gender-related vs. not) and user gender, but the criteria used to classify these dimensions are not well explained. For example, how was "gender-related aggression" operationalized? Was it pre-tested for perceived offensiveness or relevance? Please, elaborate on how the independent variables were validated and whether manipulation checks were performed.
4. There is no mention of a manipulation check to confirm that participants perceived the aggressiveness as gender-related or neutral, which is critical for internal validity in experimental designs. I suggest the authors adding or reporting results of a manipulation check to verify that participants interpreted the vignettes as intended.
5. The statistical strategy includes t-tests, Mann-Whitney U, chi-square, odds ratios, correlations, and ANCOVAs, but the rationale for each test is not always clearly articulated. It's also not clear how multiple testing and Type I error were controlled, especially with multiple group comparisons and interaction terms. The authors should clarify the logic behind test selection and whether any corrections (e.g., Bonferroni) were applied to adjust for multiple comparisons.
6. The method does not explain whether there were missing data in the surveys (other than complete non-responders) or how they were treated. Please, add details about data cleaning, imputation (if any), and handling of incomplete responses.

If applicable, is the statistical analysis and its interpretation appropriate?

1. While statistical significance is reported for several comparisons (e.g., differences between experimental and control groups, gender comparisons, ANCOVAs), the effect sizes (e.g., Cohen's d, η^2, or ηp^2) are mostly missing or inconsistently reported. This is because effect sizes are essential for understanding the practical significance of findings beyond p-values. Without them, it's hard to assess how meaningful the observed differences are.
2. I observe inconsistency in reporting total sample size where the sample size indicated in results mismatch with what the authors said in the methods section.
3. Another issue is an inconsistency in the reported sample size. While the methods section states that the total sample consists of 110 participants, including 87 female participants,

Table 2 presents a total of 119 participants

Are the conclusions drawn adequately supported by the results?
One important area for improvement in this study is the absence of key structural components that are essential for a complete and coherent scientific manuscript. Notably, the study lacks a clearly defined section on “Strengths and Limitations”—while some limitations are mentioned in the discussion, they are fragmented and not critically examined, and no strengths are explicitly highlighted. Combining strengths and limitations under one subheading would provide a balanced assessment of the study’s methodological rigor and limitations, such as the experimental vignette design, sampling issues, and the lack of cultural diversity. Additionally, the manuscript omits a dedicated “Implications” section, which is crucial for translating research findings into clinical, policy, or training recommendations—especially given the study’s focus on bias in mental health practice. Furthermore, the absence of a clear “Conclusion and future directions” section weakens the closing impact of the paper. This section is important to summarize key findings, reinforce the study’s contributions, and propose specific avenues for future research, such as exploring gender-related bias in more diverse and ecologically valid settings. Including these missing sections would significantly enhance the clarity, scientific value, and practical relevance of the study.

Is the work clearly and accurately presented and does it engage with the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results?
No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Psychology and rehabilitation sciences

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 17 March 2025

https://doi.org/10.21956/openreseurope.20921.r51165

© 2025 Fuss J et al. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Johannes Fuss
University of Duisburg-Essen, Essen, Germany
Maria Krasnova
Institute of Forensic Psychiatry and Sex Research, Center for Translational Neuro- and Behavioral Sciences, University of Duisburg-Essen, Essen, Germany
Bastian Nau
Institute of Forensic Psychiatry and Sex Research, Center for Translational Neuro- and Behavioral Sciences, University of Duisburg-Essen, Essen, Germany

Is the work clearly and accurately presented and does it cite the current literature?
The research question is timely and relevant. Gender-related differences in aggression may influence how individuals receive psychiatric care. However, I recommend reviewing more recent literature, particularly regarding the prevalence of psychiatric disorders. The cited studies on gender differences in prevalence are primarily from 1997 to 2006, whereas more recent data would provide a more comprehensive and up-to-date perspective. The manuscript also contains further statements on gender differences, which are exclusively sourced from studies conducted between the 1960s and 1990s. Given the significant expansion of research on gender and mental health in recent years, incorporating recent findings would strengthen the study's foundation and support the research focus. Additionally, some language choices in the introduction, such as the term “deranged,” may carry unintended connotations and should be revised to ensure precise and non-stigmatizing terminology.

Is the study design appropriate and does the work have academic merit?
A vignette study is generally an appropriate method for addressing this research question. However, a major concern lies in the operationalization of the vignettes. The justification for attributing verbal aggression to gender-related factors is primarily based on the phrases "the family portrayed him as a very traditional man" and "the family described her as a very modern woman." This alone does not establish a clear gender-related motivation for verbal aggression. While the assumption that a "traditional man" supports patriarchal structures and perceives himself as superior to women is somewhat understandable, the concept of a "modern woman" does not inherently imply a dominance over men. It should have been empirically tested whether participants indeed interpret verbal aggression from "traditional men" as gender-motivated and whether the same assumption holds for "modern women." The gendered nature of the aggressive comments themselves appears to play an important role, yet this aspect is not sufficiently highlighted in the article. Explicitly addressing this in the methods section would improve clarity. The sample composition presents a significant limitation. Given that gender differences are central to the study question, the considerable disparity in group sizes makes interpretation difficult and limits the validity of the findings. The underrepresentation of male participants is particularly problematic, as the authors themselves acknowledge that the limited size of the male sample has resulted in insufficient statistical power for gender-disaggregated analyses. While the authors attribute this to recruitment challenges during the COVID-19 pandemic, it is worth noting that, as a Qualtrics-based study, they could have expanded to online recruitment.

Presenting participants only with a vignette featuring a service user of the opposite sex leaves open the question whether the observed effect occurs exclusively in cross-gender interactions. It would have been methodologically beneficial to include conditions where the service user's gender matched that of the participant to explore potential within-gender effects.
Another methodological issue concerns the symptomatology presented in the vignettes. The specific symptoms included and their intended diagnostic implications are not clearly described. After reviewing the vignette, it appears that the authors aimed to depict psychotic symptoms. This should be explicitly stated in the methods section, as symptom-related variables are later analyzed in the results. The nominal categorization of symptoms into four diagnostic options seems insufficient for drawing meaningful conclusions. A more detailed diagnostic tool would have strengthened the study's validity.
The inclusion of two items from the World Values Survey lacks a clear rationale. The selection of these items is not explained, and at least one item—"A child needs a mother to grow up happy"—does not appear to have a direct connection to the study's research question. Providing a justification for their inclusion would enhance the comprehensibility of the study design.

If applicable, is the statistical analysis and its interpretation appropriate?

The rationale for conducting a correlation analysis is not clearly explained. The analysis was performed across all participants and vignettes without distinguishing between gender-related and non-gender-related vignettes or between male and female participants. This lack of differentiation limits the interpretability of the results. Additionally, multiple statistical tests were conducted, yet no correction for multiple testing was applied, increasing the likelihood of Type I errors.
Another issue is an inconsistency in the reported sample size. While the methods section states that the total sample consists of 110 participants, including 87 female participants, Table 2 presents a total of 119 participants.

Are the conclusions drawn adequately supported by the results?

Several arguments presented in the discussion to explain gender differences are difficult to follow. For instance, the statement, "Firstly, previous studies have consistently found that males tend to engage in riskier behaviors compared to females (e.g., Pawlowski et al., 2008)," is not clearly relevant, as the study under review does not investigate risky behavior. It remains unclear how this reference supports the findings.
Interpreting the results is challenging due to the small sample sizes, particularly among male participants (10 and 13 per condition). This limits the reliability of all group comparisons. Furthermore, the authors discovered significant differences in attitudes and beliefs in the coercion and paternalism scores of the BAMHS between groups. It is possible that these differences may have influenced the results; however, this potential influence is not discussed in the present study.
The claim that "the prevalence of specific diagnoses varied by gender in our study" is difficult to support given the methodological approach. Diagnosis was assessed using a single nominal question, which is not sufficient for estimating prevalence.
The interpretation of gender-related effects in the study is compromised by the study design. Female participants were only presented with male service users, and male participants only with female service users. It is difficult to determine whether any observed effects are driven by the gender of the participant, the gender of the service user, or an interaction between the two. This

significantly weakens the validity of the conclusions drawn from these findings. Furthermore, the conclusions regarding gender-related aggression differences are not necessarily supported by the results of this study. This is due to the uncertainty regarding the effectiveness of the vignettes in inducing gender-related aggression through the inclusion of a 'traditional man' and a 'very modern woman' information element. In addition, the comparability of the effectiveness of this manipulation between the two groups remains to be established.

Is the work clearly and accurately presented and does it engage with the current literature?
Partly

Is the study design appropriate and is the work technically sound?
No

Are sufficient details of methods and analysis provided to allow replication by others?
Not applicable

Are all the source data and materials underlying the results available?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Psychiatry, psychology

We confirm that we have read this submission and believe that we have an appropriate level of expertise to state that we do not consider it to be of an acceptable scientific standard, for reasons outlined above.