Gender Biases in Estimation of Others’ Pain

Lanlan Zhang,* Elizabeth A. Reynolds Losin,† Yoni K. Ashar,* Leonie Koban, † and Tor D. Wager¹

School of Leisure Sport and Management, Guangzhou Sport University, Guangzhou, China, †Department of Psychology, University of Miami, Miami, Florida, †Department of Psychiatry, Weill Cornell Medical College, New York, New York, †Paris Brain Institute, Paris, France, †Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire

Abstract: Caregiving and other interpersonal interactions often require accurate perception of others' pain from nonverbal cues, but perceivers may be subject to systematic biases based on gender, race, and other contextual factors. Such biases could contribute to systematic under-recognition and undertreatment of pain. In 2 experiments, we studied the impact of perceived patient sex on lay perceivers' pain estimates and treatment recommendations. In Experiment 1 (N = 50), perceivers viewed facial video clips of female and male patients in chronic shoulder pain and estimated patients' pain intensity. Multi-level linear modeling revealed that perceivers under-estimated female patients' pain compared with male patients, after controlling for patients' self-reported pain and pain facial expressiveness. Experiment 2 (N = 200) replicated these findings, and additionally found that 1) perceivers' pain-related gender stereotypes, specifically beliefs about typical women’s vs. men’s willingness to express pain, predicted pain estimation biases; and 2) perceivers judged female patients as relatively more likely to benefit from psychotherapy, whereas male patients were judged to benefit more from pain medicine. In both experiments, the gender bias effect size was on average 2.45 points on a 0–100 pain scale. Gender biases in pain estimation may be an obstacle to effective pain care, and experimental approaches to characterizing biases, such as the one we tested here, could inform the development of interventions to reduce such biases.

Perspective: This study identifies a bias towards underestimation of pain in female patients, which is related to gender stereotypes. The findings suggest caregivers’ or even clinicians’ pain stereotypes are a potential target for intervention.

Accurate estimation of others’ pain from nonverbal cues is an essential aspect of interpersonal communication. It forms part of the foundation for empathy and caregiving. With the high pain prevalence in the United States and worldwide, recognizing others’ pain is an increasingly valuable interpersonal skill for both clinicians and laypersons. Though pain can often be assessed through pain sufferers’ self-reports, which serve as the current “gold standard” for pain assessment in clinical contexts, recognizing pain facial expressions is an integral part of pain assessment as well.

As pain expressions are communicative behaviors, observers’ interpretations of those expressions are a crucial aspect of pain communication. Those interpretations are affected not only by the characteristics of pain expressions, but by observers’ knowledge and biases about pain, and the characteristics of pain sufferers.

The perceived sex of pain sufferers has been found to be particularly influential on estimation of others’ pain.
Biases based on perceived sex, operationalized for the present study as the tendency to perceive women as in less or more pain than men for a given level of pain self-report and expressiveness, is particularly important for understanding disparities in clinical and lay settings in terms of pain assessment and treatment. A number of studies have found that providers are more likely to recommend psychological treatment for females than males, and analgesics more frequently for males than females. Female patients also wait longer to receive analgesic medication and are less likely to receive opioid analgesics. Importantly, however, there are some studies on sex biases in pain treatment that show the opposite pattern or no sex differences. For example, Raftery et al. found that female emergency room patients received more pain medication and stronger analgesics than male patients. However, in this study increased treatment of female patients is likely because female patients reported more pain and were perceived as experiencing more pain than male patients.

Despite clinical evidence of underestimation and undertreatment of female patients' pain, laboratory findings on sex biases in pain estimation have been inconsistent. In some studies, females have been judged as being in more pain than males based on their facial expressions. In other studies, participants have judged female patients to be in less pain than males, but as higher in catastrophizing and exaggeration of pain than male patients. It seems that the inconsistency in the direction of the sex bias in pain assessment stems from whether pain facial expressions are taken into account. If females are believed to be more expressive than men for a given level of pain experience, then perceivers may discount women's expressions. In addition, females may experience more pain than males in some contexts. Controlling for objective measures of pain facial expressiveness is therefore an important step to isolating perceiver bias.

Sex biases in pain estimation and treatment may arise in part from gender stereotypes about pain. Hoffman and Tarzian provided several examples of these pain stereotypes, including “Women complain more than men; women are not accurate reporters of their pain; men are more stoic so that when they do complain of pain, “it’s real”; and women are better able to tolerate pain or have better coping skills than men.” Robinson et al. formalized measurement of gender stereotypes about pain in others in the Gender Role Expectations of Pain measure (GREP). The GREP asks about perceptions of typical men's and women's sensitivity to, endurance of, and willingness to report pain. Studies using the GREP have reported that women are viewed as more willing to report pain, more sensitive to pain and less able to endure pain than men. These gender pain stereotypes may represent mechanisms underlying gender biases in pain assessment and treatment.

Here, we attempt to resolve conflicting findings over gender biases in pain estimation and understand their underlying mechanisms by extending previous work in three critical respects. First, we compared the differences in estimation of women’s and men’s pain at the same level of pain expressiveness by controlling for patients’ self-reported pain and pain facial expressiveness both during stimulus selection and in our analyses. This is necessary because 1) the amount of pain experienced across patients is highly variable, 2) pain facial response is one of the most salient cues perceivers use to estimate pain, and 3) patients’ expressivity can affect perceivers’ estimates and their empathic accuracy. Additionally, women are believed to be more expressive than men. Controlling both factors enabled us to verify that differences in pain estimates reflected inaccurate bias, rather than accurate estimation of true sex differences in pain experience and/or expressiveness.

Second, we investigated perceivers’ pain estimation by using real chronic pain patients. Most previous studies have used hypothetical judgments, stimuli depicting experimentally induced pain in healthy participants, or computer-generated avatars. These studies are valuable, but real chronic pain is likely to be expressed, and perceived, differently.

Critically, it also remains unclear to what extent gender stereotypes in pain are accurate (reflect the reality of a given cultural context). For example, women may be more facially expressive than men for a given level of pain in some cultures and contexts. If so, it would be adaptive to use sex and gender information to bias estimates of pain, so as to achieve a more accurate pain estimate. Alternatively, bias could arise from cultural stereotypes that do not reflect people’s actual experiences. To investigate this, we tested for perceived sex bias in “pain prediction errors” — the difference between self-reported pain and perceived pain. A perceived sex bias in pain prediction errors would indicate a systematic inaccuracy in pain estimates due to perceived sex in this context.

We also extended these analyses to investigate perceived sex biases in lay perceptions of pain treatment recommendations. While there is substantial evidence of perceived sex and gender biases in pain assessment and treatment in clinical settings, fewer studies have investigated these phenomena in lay perceivers. Yet, much of the management of everyday pain complaints and clinical pain conditions takes place in home settings or community-led, group-based self-management programs for chronic pain conditions. Thus, understanding biases in pain assessment and treatment recommendations by nonmedical perceivers will provide needed additional information on the pervasiveness and mechanisms of biases outside of clinical settings.

**Experiment 1**

**Objectives and Hypotheses**

The objective of Experiment 1 was to test whether patients’ perceived sex affects perceivers’ pain estimation. We hypothesized that female patients would be
perceived to be in more pain than male patients without controlling for pain facial expression intensity and patients’ self-reported pain, but that male patients would be perceived to be in more pain than female patients at equivalent levels of pain facial expression intensity and patients’ self-reported pain.

Methods

Participants

Fifty volunteers (30 females, 20 males, mean age = 22.3 years, age range 18–53 years) recruited from Boulder, Colorado participated in the study. All participants reported having no diagnosed psychiatric or neurologic disorders and no chronic pain conditions. The study was approved by the Institutional Review Board at the University of Colorado Boulder. All participants gave written informed consent and were paid for their time.

Sex and Gender

Gender has been recently defined as the “socially constructed roles, behaviors, expressions and identities of … diverse people.” It is “usually conceptualized as a binary (girl/woman and boy/man) yet there is considerable diversity in how individuals and groups understand, experience, and express it.” Sex refers to a “set of biological attributes … primarily associated with physical and physiological features including chromosomes, gene expression, hormone levels and function, and reproductive/sexual anatomy.” While sex and gender are distinct, “perceived sex” and “perceived gender” cannot be meaningfully distinguished with the stimuli used in this study. Participants inferred both sex and gender together from observable characteristics (eg, facial and body morphology, hair, clothing, and makeup) in brief videos, and were provided no direct information about sex or gender of target patients. We have opted to use the term “perceived sex” because it perhaps more closely captures the perception that a target patient is male or female. We use the term “gender bias” because it is most prevalent in recent literature referring to disparities in treatment of males and females and because discrepancies in perceived pain and other aspects of health are likely to reflect “socially constructed roles, behaviors, expressions and identities” that “influence how people perceive themselves and each other, how they act and interact, and the distribution of power and resources in society”, in accord with recent definitions of gender.

Stimuli

Thirty-six facial video clips (3 each for 6 male patients and 6 female patients) were selected from the UNBC-McMaster Shoulder Pain Expression Archive Database (Fig 1). The database contains 200 videos of 25 shoulder pain patients captured while they were performing a series of range-of-motion tests to their affected and unaffected limbs. Each test was recorded on digital videotape with the camera focused on the face.

The database further contains information about facial pain expressiveness. In brief, each video frame was action unit (AU) coded by certified Facial Action Coding System (FACS) coders, and composite facial pain expression scores for each frame were calculated according to the Prkachin and Solomon Pain Intensity (PSPI) formula, which combined scores from four actions: brow lowering (AU4), orbital tightening (AU6 and AU7), levator contraction (AU9 and AU10) and eye closure (AU43). These four actions are thought to show sufficient consistency to be considered a “core” expression of pain. We used peak PSPI score in each video as the metric of the pain facial expressiveness for that video.

In addition to FACS scores, the UNBC-McMaster database includes patients’ self-reported pain during each sequence based on a visual analog scale (VAS) ranging from 0 to 10 (from “No pain” to “Pain as bad as could be”) and ratings from independent observers ranging from 0 to 5 (from “No pain” to “Strong pain”). We chose three video clips for each patient, one at each of three intensity levels (low, medium, or high). We determined video intensity level categorizations based on the combination of patient self-reported pain and observer ratings provided by the database (for details, please see Table 1). The number of patient exemplars was insufficient to model a random effect for patient (see Study 2 for further evidence for generalization across exemplars). Here, we averaged across the three videos of each patient to ensure independence of observations.

Video sequences were edited to 6-second duration centered approximately at the frame with the highest PSPI so as to show both the neutral expressions before and immediately after shoulder movement and the intense pain expressions during movement.

Procedure

Perceivers completed 36 trials (12 targets with 3 videos each, presented in fully randomized order). In each
trial, a fixation cross was presented first for 3 to 8 seconds. Then a 6s video clip was displayed. After a 2 to 6 seconds delay, perceivers were asked to rate how much pain the patients were experiencing, using a horizontal visual analogue scale (VAS). The scale was anchored at 0, labeled as “absolutely no pain” and 100, labeled as “worst pain possible”.

Analysis

The data were analyzed with a multilevel general linear model (GLM) testing both within- and between-subject effects, implemented in a freely available toolbox (glmmfit_multilevel function in CanlabCore; https://github.com/canlab/CanlabCore). We first estimated one model for each subject across trials. The dependent variable in the within-subject model was perceivers’ pain intensity estimates. Pain estimates of the 3 trials from each patient were averaged to ensure independence of observations, yielding 12 observations per perceiver. These within-subject models included 3 predictors: patient perceived sex (male or female), pain facial expressiveness (continuous), and patient self-reported pain (continuous). Parameter estimates from the within-subject models were then submitted to a between-subject model including an intercept term, which tested for reliable effects of these predictors across participants, and a term for participant self-reported gender (male or female) as a moderator.

To estimate effect sizes, we examined “pain prediction errors” — the difference between perceivers’ pain estimates and patients’ self-reported pain, controlling for pain facial expressiveness. To operationalize this, we repeated the above analyses but without a term for perceived patient sex. That is, the within-subject models predicted perceived patient pain intensity estimates from self-reported pain and facial expressiveness. We then extracted the model residuals (the pain prediction errors) and tested for effects of perceived patient sex on these residuals.

Results

Gender Bias in Pain Estimation

Female and male patients were not perceived to be in differing amounts of pain (female mean = 34.10, male mean = 33.39) before controlling for patient pain facial expressiveness and self-report pain, t(49) = 1.53, P = .13. This does not support the hypothesis of greater perceived pain in females overall. However, potential biases are conflated with variation in true pain and expressiveness. Pain estimates (without controlling for other factors) under each intensity level are shown in Table S2. Consistent with our hypothesis about the effects of controlling for both pain facial expressiveness (PSSI) and self-reported pain, female patients were perceived to be in less pain than male patients, t(49) = 4.11, P < .001, when controlling for these 2 variables. Female patients were perceived to be in less pain than male patients at the same level of pain expressiveness and pain self-report. As expected, increased pain facial expressiveness predicted significantly increased perceiver pain estimates, t(49) = 8.43, P < .001. In this model, we did not find a significant effect of patient self-reported pain on perceiver pain estimates, t(49) = 1.7, P = .1. This is likely because pain facial expressiveness and patient self-reported pain were correlated (within-person r = .44, P < .05), and variance related to self-reported pain was likely largely explained by facial expressiveness. Finally, perceivers’ sex did not moderate their estimation of patients’ pain, t(49) = −0.19, P = .85, that is, male and female perceivers did not differ in the estimates of male and female patients’ pain.

To estimate the gender bias effect size, we tested for effects of perceived sex on “pain prediction errors” — the difference between estimated and self-reported pain, operationalized as model residuals. We found a 2.23-point difference between male and female patients, on a 0–100 pain scale (Fig 1), with male patients’ pain overestimated female patients’ pain underestimated.

Experiment 2

Objectives and Hypotheses

In Experiment 2 we had 3 objectives: 1) to replicate gender biases observed in Experiment 1 with an enlarged patient and perceiver sample, 2) to examine whether perceived patient sex would also impact treatment recommendations regarding medication and psychotherapy, and 3) to examine whether pain-related gender stereotypes helped to explain biases in pain estimation and treatment.

We predicted that 1) male patients would be estimated to be in more pain than female patients when controlling for pain facial expressiveness and patient self-reported pain (replicating Experiment 1); 2) male patients would be prescribed greater doses of pain medicine, whereas female patients would be prescribed more sessions of psychotherapy; 27 3) pain medicine would be seen as more helpful for male patients whereas psychotherapy would be seen as more helpful for female patients; 4) the gender biases in pain estimation and pain treatment would be influenced by the pain-specific gender stereotypes perceivers held.

Methods

Participants

Two hundred participants were tested via Amazon Mechanical Turk using Qualtrics software (Qualtrics, Provo, UT. http://www.qualtrics.com). More than half (108 of 200) of them have or had at least one chronic pain condition such as back pain, arthritis, migraine, etc., or acute pain conditions such as accidents, injuries, surgeries, etc. Only 3 out of the 200 participants had worked in healthcare. One participant failed to follow the study instructions (giving the same rating to all pain video stimuli) and was therefore excluded. Two
participants chose their gender identity as “other.” Given that there were only 2 participants with non-binary gender, we chose to exclude them from the analysis as it would not have been possible to conduct a meaningful test of whether this group differed from those identifying as male and female in our sample. Data from the remaining 197 perceivers (81 females, 116 males, mean age = 32.3, age range 19–55 years old) were included in the analyses. The study was approved by the Institutional Review Board at the University of Colorado Boulder. Online informed consent was provided by all participants.

**Stimuli**

Two different video clips of the facial expressions of each of 25 (13 females, 12 males) patients from the whole UNBC-McMaster Shoulder Pain Expression Archive Database were selected and edited to 6 seconds as in Experiment 1. We divided stimuli into 2 stimulus sets, with one video of each patient randomly assigned to each set. We controlled for any potential differences between the two stimulus sets by controlling for pain facial expressiveness and patient pain self-report in all of our models. We used this approach to reduce the influence of idiosyncrasies in individual patient videos while ensuring independence of observations. Each stimulus set was tested with 100 perceivers, yielding N = 98 (43 female) for stimulus set 1 and N = 99 (38 female) for stimulus set 2 (total N = 197) analyzed for each stimulus set, with random assignment of participants to a stimulus set.

**Procedure**

The online survey opened with the consent form and a demographic questionnaire. Experiment instructions (see Supplementary Materials) and a practice trial was provided prior to the main task. In each trial of the main task, perceivers first watched a 6s video clip and then estimated the patient’s pain on a horizontal VAS from 0 (“absolutely no pain”) to 100 (“worst pain possible”), as in Experiment 1. The videos were presented in a unique random order to each perceiver.

Following the main task, perceivers were then asked to prescribe pain treatments for each patient as if they were doctors (see Supplementary Materials for detailed instructions). A still frame from the video showing neutral facial expression was presented along with 3 prescription questions: 1) “If you were to prescribe pain medicine, what dose would you prescribe to this patient?”; 2) “If you were to prescribe psychotherapy, how many sessions would you prescribe?”; and 3) “What do you think would help the patient more, pain medicine or psychotherapy?” Since we believed most perceivers would not have medical training (confirmed in our sample with only 3 of 200 with health care work experience), response ranges for the first two questions were made on a VAS from 0 (minimal dose or sessions) to 100 (maximal dose or sessions), rather than using a more specific clinical unit. For the third question, perceivers were required to make a forced choice between pain medicine and psychotherapy.

Lastly, perceivers were asked to complete the modified Gender Role Expectation of Pain Questionnaire (GREP), which measures gender-related stereotypic attributions about sensitivity to pain, endurance of pain, and willingness to report pain. The 9 questionnaire items were: “What is the typical woman’s sensitivity to pain/endurance of pain/willingness to report pain?”; “What is the typical man’s sensitivity to pain/endurance of pain/willingness to report pain?”; and “What is your sensitivity to pain/endurance of pain/willingness to report pain?”.

**Analysis**

We ran three sets of analyses corresponding to our three objectives. The first set of analyses tested for gender bias in pain estimation using a multilevel GLM, as in Experiment 1.

The second set of analyses tested for gender biases in pain treatment recommendations using two analyses. First, we tested for biases in treatment recommendations (doses of medicine or sessions of psychotherapy) using separate multilevel GLMs, as in Experiment 1. Second, we tested for biases in treatment preference by calculating the proportion of the 25 trials in which medicine versus psychotherapy was prescribed across both male and female patients, and to male and female patients separately. To test overall treatment preference, we conducted a one-sample t-test against a null hypothesis of .5, which would indicate equivalent preferences for the 2 treatments. To test for gender biases in treatment preference, we performed paired t-tests on proportions for males versus females.

The third set of analyses tested whether observed gender biases in pain estimation and treatment recommendations were moderated by perceiver sex and perceivers’ pain-related gender role expectations. We used a 2-step process in which we first screened for pain-related gender role expectations that showed some relationship with pain estimation or treatment recommendations by computing Pearson’s correlations between pain-related gender role expectations and each of the 2 variables of interest: 1) participants’ average pain estimation for female and male patients, and 2) hypothetical prescriptions of pain treatment recommendations for female patients and male patients. For the second step, we included any pain-related gender role expectations showing significant correlations with pain estimation and/or treatment variables in the
multilevel GLM as moderators in the form of person-level gender role stereotype scores. The stereotype scores were created for each item by subtracting each perceiver's beliefs about typical women from those for typical men, resulting in 3 scores ranging from −100 (strongest bias in favor of thinking the typical man has higher pain sensitivity/ endurance/ or willingness to report than the typical woman) to 100 (strongest bias in favor of thinking the typical woman has higher pain sensitivity/ endurance/ or willingness than the typical man).

Results

Gender Bias in Pain Estimation

Female patients were perceived to be in less pain than male patients before controlling for other factors, t (196) = 3.27, P < .001. This is counter to the hypothesis that females would be perceived to be in more pain but consistent with our findings in Experiment 1 after controlling for pain facial expressiveness and self-reported pain. As in Experiment 1, increased pain facial expressiveness predicted significantly increased pain estimates when controlling for self-reported pain, t(196) = 19.98, P < .001. Increased self-reported pain also independently, significantly predicted increased pain estimates controlling for pain facial expressiveness, t(196) = 6.04, P < .001. More importantly, when controlling for both pain facial expressiveness and patient self-reported pain, female patients were estimated to be in less pain than male patients, consistent with our hypothesis and the findings of Experiment 1, t(196) = 4.13, P < .001. As in Experiment 1, female patients were perceived to be in less pain than male patients at the same level of pain expressiveness and pain self-report. This bias was not moderated by perceivers' gender, t(196) = −0.3, P = .76, indicating that male and female perceivers did not differ in pain estimation biases.

As in Experiment 1, we also tested the influence of perceived sex on pain prediction errors. We found more positive pain prediction errors for men (pain overestimation) and relatively more negative prediction errors for women (pain underestimation). Perceived sex introduced a bias of 2.67 points on a 0−100 pain scale (Fig 2).

Gender Bias in Pain Treatment Recommendations

Contrary to our hypotheses, the dose of pain medicine prescribed to female and male patients did not differ after controlling for pain facial expressiveness and patient self-reported pain, t(196) = .13, P = .9. As expected, however, increased pain facial expressiveness significantly predicted increased doses of pain medicine, t(196) = 4.35, P < .001. Patient self-reported pain did not predict dose of pain medicine after controlling for expressiveness, t(196) = .18, P = .86.

Also contrary to our hypotheses, the sessions of psychotherapy prescribed for female and male patients did not differ, t(196) = −.46, P = .64, when controlling for pain facial expressiveness and patient self-reported pain. As with pain medicine, however, increased pain facial expressiveness significantly predicted prescription of more sessions of psychotherapy, t(196) = 2.67, P < .01. We did not find an effect of patient self-reported pain on sessions of psychotherapy prescribed, t(196) = −.71, P = .48.

Male and female perceivers also did not differ in the doses of pain medicine (P = .62) or sessions of psychotherapy they prescribed (P = .26). Importantly, the fact that patients' pain facial expressiveness significantly predicted the dose of pain medicine and the sessions of psychotherapy prescribed suggested perceivers did not prescribe randomly, but likely based their management recommendations on their memory of patients' pain expressiveness.

When perceivers were asked to make a forced-choice recommendation about which treatment would be more helpful for each patient, pain medicine was preferred over psychotherapy overall for both females (58% medicine vs 42% psychotherapy) and males (62% medicine vs 38% psychotherapy). The average number of trials for which each participant prescribed pain medicine was significantly greater than 50%, t(196) = −5.59, P < .001. A comparison of treatment preference for male and female patients revealed that psychotherapy was preferred for a higher proportion of female patients (42%) than male patients (38%). A paired t-test on the average number of trials for which each participant prescribed psychotherapy for females vs. males showed a significant difference; t(196) = 2.47, P < .05 (Fig 3).

Moderation of Bias in Pain Estimation and Pain Treatment Recommendations by Pain-Related Gender Role Expectations

We found a significant interaction between perceiver gender and typical gender judgments (F(1, 195) = 9.08, P < .01) showing that female perceivers believed that
Figure 3. Pain treatment preference for psychotherapy vs. medication for male and female patients. Y axis values represent the proportion of the patients of a given gender in which perceivers indicated psychotherapy would be more helpful than pain medication in a forced-choice test. Proportion of pain medicine = 1 - proportion of psychotherapy. Psychotherapy was preferred more often for female than for male patients whereas pain medication was preferred more often for male than female patients. Error bars reflect the standard error of the mean. *, P < .05.

Table 1. Correlation Between Pain-Specific Gender Role Expectations and Pain Estimation or Treatment Recommendations

<table>
<thead>
<tr>
<th>GREP</th>
<th>Pain Estimation</th>
<th>Pain Medicine</th>
<th>Psychotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>TypWoEnd</td>
<td>0.05</td>
<td>-0.2</td>
<td>0.01</td>
</tr>
<tr>
<td>TypMeEnd</td>
<td>-0.16*</td>
<td>0.18*</td>
<td>0.12</td>
</tr>
<tr>
<td>TypWoWil</td>
<td>0.11</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>TypMeWil</td>
<td>0.14*</td>
<td>0.21**</td>
<td>0.29***</td>
</tr>
</tbody>
</table>

GREP, Gender Role Expectation of Pain Questionnaire. TypWoEnd, typical woman’s willingness to report pain; TypMeEnd, typical man’s endurance of pain; TypWoWil, typical woman’s willingness to perform pain; TypMeWil, typical man’s willingness to report pain. FePat, female patients; MaPat, male patients; *P < .1; **P < .05; ***P < .01; ****P < .001.

Discussion

Accurate estimation of others’ pain is an important interpersonal skill in both clinical and everyday settings. We found that lay perceivers’ pain estimation and pain treatment recommendations were biased by perceived patient sex. Perceivers estimated female patients to be in less pain than male patients at the same level of pain expressiveness and pain self-report. They underestimated women’s pain relative to actual self-reported pain, while overestimating men’s pain relative to self-report. Perceivers also indicated that female patients would benefit more than male patients from psychotherapy vs. medication when choosing between the two. Furthermore, we found that perceivers’ pain-related gender stereotypes influenced their biases in pain estimation. The more willing perceivers believed women are report pain than men, the less pain they perceived female patients to be in less pain than men, t (196) = -2.45, P < .05. In contrast, gender stereotypes about endurance of pain did not moderate gender biases in pain estimation, t(196) = .75, P = .46. Importantly, female patients were still perceived to be in less pain than male patients after adding gender stereotypes about willingness to report pain to the model, t (196) = 4.43, P < .001, indicating that the average participant shows reasonably strong gender bias and that other factors such as additional pain stereotypes, or perceptual biases likely contribute to this bias. For pain treatment recommendations, neither gender stereotype moderated the gender bias in hypothetical prescriptions of medicine or psychotherapy.
estimation and treatment biases. These findings add to a growing body of evidence of demographic biases in the estimation and treatment of others’ suffering in a wide range of societal domains from the medical field\(^{26,27,56,65}\) to the justice system\(^{41}\) and point to stereotype education as a promising avenue for mitigating these biases.

**Perceived Sex Effects on Pain Estimation and Treatment Recommendations**

Male patients were perceived to be in more pain than female patients both before and after controlling for pain facial expressions and pain self-report in Experiment 2. This same pattern was only revealed after controlling for pain facial expressions and pain self-report in Experiment 1, with no perceived sex difference in pain estimates evident before controlling for these factors. The finding that men’s pain is overestimated relative to women’s is consistent with other studies suggesting that males’ pain is perceived as more intense than female’s pain,\(^{56}\) even in neonates.\(^{10}\) However, our finding stands in contrast to an even larger number of studies that have found increased estimation women’s pain compared to that of men.\(^{26,36,51}\) In addition, a substantial literature demonstrates that when sex or gender differences in pain report are present, women are almost always found to report more pain than men.\(^{16,20}\)

Several factors may explain these discrepancies, including variability across studies in whether perceivers judge specific exemplars (eg, individuals in pain or facial expressions) vs. make normative judgments about pain in general, and the specific aspects of pain that are being judged. For example, perceivers may simultaneously believe that women experience more pain in response to a given stimulus, but undervalue the actual indicators of pain in women. Our perceivers rated genuine pain expressions in clinical shoulder pain patients. This, the gender biases observed in the present study may be more similar to those seen in real-world clinical and lay pain evaluation contexts than in some previous studies.

Another factor that may help explain differences between our findings of biases in pain estimation and those in previous studies is whether pain expressiveness and self-report are controlled for. Women are consistently found to report higher levels of pain than men\(^{6,20}\) and to be more expressive of pain than men.\(^{22,60}\) Perceivers may in turn get habituated to more frequent or more intense pain expressions in females and as a result reduce the pain they attribute to those expressions. A similar bias has been seen upon repeated exposure to pain facial expressions, which has been found to reduce vicarious pain estimation.\(^{21,46}\) Thus, controlling for both private self-report and observable cues is critical to test for the presence of a true bias, that is, differences in pain estimates for male and female patients based on the same pain cues. Several previous studies in which women were perceived to be in more pain than men did not control for pain facial expression intensity, and to our knowledge, no prior studies controlled for the pain sufferer’s own pain reports, either alone or in combination with their pain facial expressions. Thus, some of those prior findings may not have reflected true biases in pain estimation. Furthermore, our findings of men being perceived in higher pain than women are consistent with biases in clinical pain treatment, which have demonstrated that men are more likely to receive analgesic treatment, in stronger forms, and more rapidly than women.\(^{7,24,27}\)

A final factor that may contribute to the difference between our findings and those in previous studies is that some previous studies employed observers with medical training. In contrast, only 3 of 200 participants in Experiment 2 in the present study reported having healthcare work experience. Previous studies have suggested that medical training may influence how people perceive pain in others. For example, health professionals have been found to underestimate patients’ pain.\(^{58}\) Although we do not know of any studies showing that medical training changes pain assessment and treatment biases related to patient demographics, this is also plausible.

**Gender-Related Pain Stereotypes Contribute to Perceived Sex Differences in Pain Estimation and Treatment Recommendations**

Prior literature suggests the most likely source of gender biases in pain estimation is the norms and stereotypes perceivers hold about emotional expression and pain responses. Generally, boys are discouraged from expressing emotions, whereas girls are permitted to express them. These norms hold in many cultures and societies from a young age.\(^{28,43}\) As a result, men may be more reluctant to express pain and other vulnerabilities than women.\(^{43}\) Thus, masculine gender norms are associated with high pain tolerance and stoicism whereas feminine gender norms are more permissive of expressing pain.\(^{39,61}\)

We found evidence of these pain-related gender stereotypes in the current study. Specifically, women were considered substantially more willing to report pain than men by both male and female perceivers, replicating prior work.\(^{50,53,66}\) Additionally, female perceivers rated women as more willing to endure pain than men. Critically, individual differences in the beliefs regarding women’s higher willingness to report pain were directly linked to reduced pain estimates of female (compared to male) patients. Thus, it is plausible that male patients’ expressions of pain were seen as more credible whereas those of female patients were discounted during pain estimation, to correct for an assumed higher pain expressiveness of women. We found no statistically significant perceived sex differences in pain facial expressiveness in the present stimuli (see Supplementary Materials), but other studies suggest that females may indeed report and express higher pain than males.\(^{17,20,22}\) In other words, our participants’ bias (stereotype) that women are more expressive of pain than
men may well be accurate at the population level. However, it led them to misestimate pain in this study. Indeed, studies have shown that people use group-level stereotypes when information about the individual person is sparse, as was the case for our stimuli and in many medical encounters. Alternatively, perceivers may habituate to repeated and more frequent expressions of pain in females. Grégoire et al. showed that repeated exposure to others’ pain reduces vicarious pain intensity estimation. Using signal-detection techniques, Prkachin et al. similarly demonstrated that repeated exposure to pain facial expressions led to more conservative recommendations about pain estimation. Although in the present study there were no sex differences in pain facial expressiveness in video stimuli, it is possible that perceivers’ previous experiences of higher pain expressiveness by women compared to men made them less willing to attribute pain to female patients. If this was the case, our findings, together with evidence from other studies, would suggest that frequent exposure to female patients’ intense pain facial expressions in clinical settings might contribute to underestimation of their pain and in turn may result in substandard treatment for them.

In contrast to our findings on willingness to report pain, gender stereotypes about endurance of pain did not moderate perceived sex biases in pain estimation. This finding may be in conflict with the findings of Robinson and Wise who report that both male and female gender pain stereotypes predict pain estimation for male and female patients. However, our use of a difference score between female and male gender pain stereotypes as a moderator of perceived sex effects on pain estimation is not directly comparable to the analyses in Robinson and Wise. Our finding that female perceivers rated women as more willing to endure pain than men but male perceivers did not is consistent with the findings of Wandner et al. but inconsistent with the findings of Robinson et al., who found that both males and females rated women as less willing to endure pain than men. Finally, our finding that neither male nor female perceivers rated typical men and women to differ in pain sensitivity is in conflict with the findings of both Robinson et al. and Wandner and colleagues, who both found that women were perceived to be less willing to endure pain and more sensitive to pain by both male and female perceivers. One factor that may explain the differences between our findings and those of Robinson et al. is that participants viewed patient pain videos prior to filling out questionnaires about their gender pain stereotypes in our study. Patient pain videos may have influenced our participant’s gender stereotypic pain views, at least in the short term.

**Implications of Underestimation and Psychologization in Treatment of Women’s Pain**

A critical question for understanding the potential effects of gender biases in pain estimation is whether they are accurate at a population level. For example, if women in a given cultural context are more expressive relative to their level of pain compared to men, then perhaps a gender bias would be adaptive in forming accurate pain estimates at a population level. Taking self-reported pain as the ground truth, however, our analyses of pain prediction error in both experiments show that perceived sex introduces a systematic bias leading to inaccurate pain estimation, with underestimation of women’s pain and overestimation of men’s relative to their reports of their own pain.

Pain underestimation and psychologization in the treatment of women’s pain could both have detrimental effects on female patients’ health outcomes. For example, in Schäfer et al., health care providers who believed female chronic pain patients were more likely to exaggerate their pain prescribed psychotherapy to female patients but opioids to male patients, presumably because providers attributed less pain to female patients. Yet, prior studies have not found consistent differences in the treatment efficacy of psychotherapy and opioids for female and male patients. Thus, the belief that psychotherapy is more helpful to female patients and opioids are more helpful to male patients likely represents a true bias, that is, a deviation from the most efficacious treatment strategies. Our data cannot speak to the particular reason that participants were more likely to prefer psychotherapy for women, however. One possibility is that participants may believe that women’s pain is less nociceptive in origin. Another is that they may believe women will be more receptive to psychotherapy and benefit more from it. Other explanations may be possible as well.

**Limitations**

Our findings should be interpreted in light of several limitations. The size of the target patient sample used in this study is limited (N = 12 in Experiment 1 and N = 25 in Experiment 2). Because of the limited number of available stimuli, our analyses treated patient as a fixed (nonrandom) effect. Therefore, future studies will need to test whether our findings generalize to other populations of male and female patient exemplars. Additionally, because the factors affecting the estimation and treatment of patient pain may differ across different pain conditions, it is unclear whether our findings will generalize to pain disorders beyond shoulder pain. However, we did replicate the main findings across two experiments.

Another aspect that potentially limits the generalizability of the findings to clinical settings is that hardly any of the pain perceivers had medical training or healthcare work experience (3/200 in Experiment 2). However nearly half of them (92 of 200 in Experiment 2) reported past or current experiences with pain, thus they likely did have experience with considering treatments for their own pain and likely the pain of those in their families. Indeed, pain is a universal condition, and many recommendations about pain and its management are made outside of medical care by members of the lay public. Even individuals with chronic pain conditions often make
recommendations about their own pain treatment, striving for self-management. Furthermore, lay-led, group-based self-management programs for patients with chronic pain have been developed, which entail the evaluation of pain and pain treatment by lay perceivers. Therefore, understanding the biases that lay perceivers have in both pain assessment and treatment recommendations has value and potential clinical relevance.

The magnitude of the perceived sex bias in pain estimation revealed in this study was on average 2.45 points out of a 100-point scale. Further studies are needed to decide whether an effect of this size would be practically meaningful, though small-magnitude effects can have practical significance when they persist across large numbers of individuals in a society. Similarly, the correlations between perceived sex bias in pain estimation and gender role expectations in our studies, though small, are statistically significant. Also importantly, female patients were still perceived to be in less pain than male patients after adding gender stereotypes about willingness to report pain to the model.

This study was not able to address biases related to race and age (for studies, see Hollingshead et al as the UNBC-McMaster Shoulder Pain Expression Archive Database does not systematically vary in race and age, nor does it provide demographic information for the patients. As studies with virtual agents have demonstrated that the perceived age and race/ethnicity of the target can also influence pain assessment and treatment, it is possible these factors could have influenced our results. However, the patients appear to be mostly middle-aged Caucasian adults. Therefore, we think race and age are unlikely to represent a major confound in findings. Generalization of our findings to other patient demographic groups will require further research with stimuli systematically varying in other demographic dimensions.

Lastly, averaging the pain estimates of the three trials from each patient, though ensuring the independence of observation, obscures potential interactions with painful stimulus intensity that could further inform interventions aimed at mitigating gender disparities in pain treatment. Similarly, as we only included pain facial expressions, we were not able to test the specificity of the biases we observed to pain. The degree to which our findings generalize across different levels of observed pain intensity and different facial expressions should be addressed in future studies.

Conclusions

The present study demonstrates a perceived sex bias in lay perceivers’ pain estimation and pain treatment recommendations. Women’s pain was viewed as less intense and as benefiting more from psychotherapy than men’s pain, and perceivers’ gender role stereotypes about willingness to report pain explained some of this bias in pain estimation. These findings suggest that gender biases in pain estimation may be an obstacle to effective pain care, and that experimental approaches to characterizing biases, such as the one we tested here, could inform the development of interventions to reduce such biases.

Acknowledgments

The authors would like to acknowledge Daniel Kusko, Daniel Weflen and Philip Kragel for their help on participant recruitment, code for data integration and analysis, and helpful discussion.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jpain.2021.03.001.

References


12 The Journal of Pain


