Original experimental

Marie Gustafsson Sendén and Emma A. Renström*

Gender bias in assessment of future work ability among pain patients – an experimental vignette study of medical students' assessment

https://doi.org/10.1515/sjpain-2018-0108
Received June 29, 2018; revised November 8, 2018; accepted November 13, 2018; previously published online January 17, 2019

Abstract

Background and aims: Pain is a prevalent problem in many countries. Women are more often on sick-leave for pain than men. Such gender differences have been explained through biological factors, different demands for on the job market, and home conditions. Fewer studies have focused on how gender stereotypes may bias the medical assessment of pain patients. The aim of the present research was to investigate if a gender bias in medical students' evaluations of chronic pain patients can contribute to explaining the gender differences in sick-leave due to pain. Specifically, we investigated whether medical students' estimates of a patient's accuracy of their own work ability and amount of domestic work differed between female and male patients, and how such estimates influenced the medical students' judgments of the patient's work ability.
Methods: Medical students ($n = 137$; 60 women; 74 men; three unspecified) read a vignette describing a patient with pain and filled out a questionnaire. The vignette was identical and gender neutral, except for the name of the patient signaling gender. A between-subjects experimental design was used in which participants were randomly assigned to an experimental condition. Participants then judged the patient's work ability, the accuracy of the patient's self-assessed work ability, and the amount of domestic work they believed was performed by the patient. All ratings were made on seven-point items.
Results: The results showed that there was no main effect of gender on perceived future work ability of the patient, $F\ (1,131) = 0.867$, $p = 0.353$. However, male patients were considered to be more accurate in their self-assessed work ability than female patients $F\ (1,131) = 5.925$ $p = 0.016$ ($M_{female} = 4.87$, $SD_{female} = 1.22$, and $M_{male} = 5.33$, $SD_{male} = 1.02$). Moreover, female patients were thought to perform more domestic work, $F\ (1,131) = 25.56$, $p < 0.001$ ($M_{female} = 4.14$, $SD_{female} = 1.41$, and $M_{male} = 3.07$, $SD_{male} = 1.16$). Finally, perceived amount of domestic work moderated the effects of perceived future work ability for female but not for male patients, $B = 0.42$, $p = 0.005$. Hence, there was a positive effect of amount of domestic work performed on work ability judgments for women, such that the more domestic work they were assumed to perform, the more they were perceived to be able to work.
Conclusions: Gender stereotypes influenced assessments of future work ability in pain patients, mainly because women were assumed to perform more domestic work which had a positive effect on perceived work ability. Because domestic work should have a negative effect on recovery, expectations from the physician that domestic work is expected by female patients may in fact have the opposite effect prolonging sick-leave. Moreover, the students trusted the male patients' ability to assess their own work capacity more than women's.
Implications: It is important that medical students receive education about gender biases and how they may influence medical assessment during their training. Such education may alleviate the influence of gender stereotypes.

Keywords: gender bias; medical assessment; pain patient; sick-leave; return to work.

***Corresponding author: Emma A. Renström,** Department of Psychology, University of Gothenburg, Gothenburg, Sweden, E-mail: emma.renstrom@psy.gu.se
Marie Gustafsson Sendén: Department of Psychology, Stockholm University, Stockholm, Sweden; and Department of Social Sciences, Södertörn University, Stockholm, Sweden

1 Introduction

Pain is one of the most common reasons for long term sick-leave [1], and prevalence is systematically higher among women than men. Women more often seek medical care

for pain-related problems and are listed for sick leave more often and for longer periods than men [2–5].

Gender differences in pain and sick leave could be rooted in biological factors [3], differential job markets, work and home conditions [6], that make women more vulnerable to pain problems. Furthermore, gender differences might occur in diagnosis and treatment processes because of gender stereotypes [7, 8]. A stereotype is a preconception of the traits and characteristics associated with a social group [9]. Stereotypes are both descriptive and prescriptive meaning that they both reflect how reality is constituted, but also prescribe how group members of different social groups should behave [10]. Violating stereotypical roles are associated with social sanctions, while the conformity to such roles are rewarded [11, 12]. Gender stereotypes are hence mental representations about what is typical for women and men [13, 14], and affect expectations of behavior, abilities, duties and responsibilities. On average, gender stereotypes might reflect reality, but when evaluating specific individuals stereotypes might lead to the wrong decisions [9].

The traditional female stereotype is associated with domestic and caretaking roles, while the traditional male stereotype is to be the family breadwinner [15, 16]. Thus, women are associated with being at home and men are associated with being away from home, and specifically at work. Such stereotypes could influence who is prescribed sick-leave since being at home aligns more with the female stereotype than with the male stereotype.

Gender stereotypes are especially influential in ambiguous situations, and when social judgments are performed under time pressure [17]. This means, that when the full history of a target individual is not accessible, people are guided by categorizations and top down processing of information. The process of sick listing and assessments of work ability is rather ambiguous because there are no general agreements on how to define and assess work ability [18]. This leaves the concept of work ability open to personal and intuitive interpretations, leading to an enhanced risk that gender stereotypes may influence the judgement process.

Previous research has shown gender differences in diagnosis, medical contacts and treatment. For example, women's pain symptoms were more often than men's explained by psychosocial factors [19], and given non-specific diagnoses [8]. Family aspects were attended to in 61% of the female cases and 37% of the male cases, which indicates that the family situation is given priority in assessments of women. Hence, it could be hypothesized that perceptions of amount of domestic work a patient performs both is influenced by patient gender, but also in turn influence assessments of work ability, especially for female patients.

The patient's own perception of their work ability is one of the strongest predictors of decisions for sick listing, and return to work [20]. In relation to gender stereotypes, studies have found that men expressed higher demands toward their physician and experienced higher support from the physician [21], and were taken more seriously [22], than female patients [23]. Thus, physicians may attend more to male patients' self-assessment as compared to female patients' self-assessment. Hence, perceptions of the patient's accuracy in assessing their own work ability may be more influential in judgments of male patients as compared to female patients.

The present study focuses on the assessments of a fictive patient's future work ability, and how such assessments vary with patient gender. From a psychological perspective, expectations might influence the rehabilitation process substantially [24–26]. Hence, deeper insights about the mechanisms operating in sick-listing processes are important. Based on the literature review we first formulate main effects hypotheses and predict that:

1. Female patients will be ascribed lower future work ability than male patients.
2. Female patients will be perceived as doing more domestic work than male patients.
3. Male patients will be perceived as more accurate in assessing their own work ability than female patients.

Moreover, as indicated above, different factors should affect judgments of work ability differently for women and men. Hence, we also formulate two conditional hypotheses and predict that:

4. For male, but not female, patients, there will be a positive effect of patient accuracy on estimates of work ability.
5. For female, but not male, patients, there will be a positive effect of amount of domestic work performed by the patient on estimates of work ability.

When it comes to gender of the physician, previous research has shown inconclusive results [27]. Therefore, we include the gender of the participant as a control variable.

2 Methods

2.1 Design

The design was a two-factor between-subjects design where gender of a described patient was manipulated. Participants

were randomly assigned to read a description of a pain patient. The description was exactly the same in both conditions except for the name of the patient indicating that the patient was a woman or man. The names were traditional Swedish names for a woman (Karin) or a man (Mats).

2.2 Participants

Medical students at a large Swedish University hospital participated in the study ($n = 137$; 60 women, 74 men, and three that did not provide gender information). Because gender of the physician was a control variable, those who did not provide gender information were excluded from the analyses where gender of the participant was included. Two persons did not complete the questionnaire. The mean age was 23.76 (SD = 4.62). The sample size was determined by the number of predictors in the regression [28].

2.3 Materials

2.3.1 Stimuli material

The text presented to the participants described a patient with a gender-neutral pain symptom. The patient was described as suffering from back pain that started when moving to a new house. After having visited primary care the patient was now on sick leave due to the pain. Further, the patient was described as having a spouse and two children, and that the pain had affected the family relations. The text was carefully edited to be gender neutral. For instance, we did not reveal how the pain had affected the family relations, and the patient was described as having a neutral job at an office, that he/she enjoyed traveling and spending time with family and friends.

2.3.2 Future work ability assessment

The main dependent variable *Future Work ability* was measured with the item "How do you assess the patient's work ability within the next three years?". Responses were assessed on a seven-point Likert scale from *1 = very poor* to *7 = very good*.

2.3.3 Domestic work assessment

Domestic work was measured with the item "How much domestic work do you think Karin/Mats does while on sick leave?". Responses were assessed on a seven-point Likert scale from *1 = very little* to *7 = very much*.

2.3.4 Perceived accuracy of the patient

Patient's accuracy was measured with the item "(Karin/Mats) considers (his/her) own ability to work as high. How accurate do you think that judgment is?" Responses were assessed on a seven-point Likert scale from *1 = not at all accurate* to *7 = very accurate*.

2.4 Procedure

Participants were approached at campus by an experiment leader and asked to participate in a questionnaire study on medical assessments. Participants were orally informed about anonymity, right to withdrawal and data handling. Participants who agreed to participate were also given written information about informed consent before they started to fill out the questionnaire. Participation was reimbursed with an ice-cream voucher. The questionnaires were mixed before distribution and the first page was identical in both versions. Hence, the experiment leader was blind to conditions.

First, participants read a one-page description of a patient with pain symptoms, and then rated questions about the patient's work ability, patient's accuracy and patient's amount of domestic work. After this, they answered some demographical questions. Participation took about 10 min and upon returning the questionnaire to the experiment leader, participants were thanked and reimbursed. Upon handing in the questionnaire, they were also asked about suspicions of the purpose of the study. None indicated any suspicions about the gender aspects of the study.

2.5 Statistical analyses

ANCOVA's and hierarchal linear regressions analysis were performed in SPSS (IBM SPSS Statistics for Windows, Version 24.0, IBM Corp., Armonk, NY, USA) and *p*-values of 0.05 or less were considered significant.

To test hypotheses 1–3, we performed three separate ANCOVA's, one for perceived work ability, one for patient accuracy and one for amount of domestic work, including the gender of the patient as the independent variable, and the gender of the participant as a covariate.

In testing hypotheses 4 and 5, we computed a hierarchical linear regression analysis, in which the independent

variables (patient gender, patient's accuracy, and domestic work), and their interaction with patient gender were used as predictors for the dependent variable future work ability. We also controlled for participant gender, and the interaction between participant gender and patient gender. The regression was computed in two steps. In step 1, the variables patient gender (female = 1, male = 0), participant gender (female = 1, male = 0), domestic work and patient's accuracy was entered. In step 2, the bivariate interaction terms between patient gender and domestic work, patient gender and patient's accuracy in self-assessment, and patient gender and participant gender were added.

3 Results

Means and standard deviations of the study variables are presented in Table 1. The first hypothesis stated that female patients would be judged to have lower future work ability. However, the first ANCOVA showed that there was no effect of patient gender, F (1,131) = 0.867, p = 0.353, mean difference = 0.19 (M_{male}–M_{female}), 95% CI: Lower bound = –0.2; Upper bound = 0.58, indicating that the hypothesis was not supported. The second hypothesis stated that male patients would be considered more accurate in assessing their own work ability than female patients, which was supported by the analysis, F (1,131) = 5.925 p = 0.016, η_p^2 = 0.04, mean difference = 0.47 (M_{male}–M_{female}), 95% CI: Lower bound = 0.09; Upper bound = 0.86. The third hypothesis stated that female patients would be judged as doing more domestic work than men, and the results supported this, F (1,131) = 25.56, p < 0.001, η_p^2 = 0.16 mean difference = –1.13 (M_{male}–M_{female}), 95% CI: Lower bound = –1.57; Upper bound = –0.69.

Hypotheses 4 and 5 were tested using hierarchical regression analysis as described above. The results are presented in Table 2.

The first model was statistically significant, F (4,128) = 8.670, p < 0.001, R^2 = 0.201. As can be seen in Table 2, domestic work had a statistically significant positive effect on work ability assessments such that the more a patient was perceived to engage in domestic work, the more likely they were thought to be able to work within the coming years. There was also a statistically significant positive effect of patient's accuracy such that the better the patient was perceived to be at judging their own work ability, the higher the expected work ability was perceived to be. Hence, if the patient believes to be able to work, the medical students also thought that they would be able to work.

The explained variance of the model increased statistically significantly when adding the interaction terms in step 2, F (7,125) = 6.654, p < 0.001, ΔR^2 = 0.052, p = 0.029. Hypothesis 4 stated that for male, but not female, patients there would be a positive effect of patient accuracy on estimates of future work ability. However, the interaction between patient accuracy and patient gender was not statistically significant. Hypothesis 5 stated that for female, but not male, patients there would be a positive effect of amount of domestic work on estimates of future work ability. The results showed a statistically significant interaction effect between patient gender and domestic work on judgment of work ability. The interaction is plotted in Fig. 1. For female patients, the judgment of their future work ability depends on the amount of domestic work they are perceived to perform. However, for male patients, amount of domestic work has no influence on judgments of work ability. The simple slopes analysis showed that for the female patient the slope was statistically significant, B = 0.37, p < 0.001, while the slope for the male patient was not, B = –0.05, p = 0.66.

4 Discussion

The present research showed that even though male and female patients were expected to have similar future work ability, patient gender influenced the decision process. First, participants were influenced by gender when they

Table 1: Means and standard deviations for the main study variables, for female and male patients separately as well as for the whole sample.

	Female patient (n = 66)		Male patient (n = 70)		Total (n = 136)	
	Means	**Standard devations**	**Means**	**Standard deviations**	**Means**	**Standard deviations**
Future work ability	3.55	1.12	3.43	1.17	3.64	1.14
Patient's accuracy	4.87	1.22	5.33	1.02	5.10	1.14
Domestic work	4.14	1.41	3.07	1.16	3.59	1.39

All ratings were made on seven-point scales where future work ability was rated between *1 = Very poor* to *7 = Very good*, perceived accuracy of the patient's self-assessed work ability was rated between *1 = Not at all accurate* to *7 = Very accurate*, and perceived amount of domestic work was rated between *1 = Very little* to *7 = Very much*.

Table 2: Hierarchical linear regression analysis in two steps, with single predictors in step 1 and interaction terms added in step 2, with work ability as the main outcome variable.

	Step 1 ($r^2 = 0.201$)			Step 2 ($r^2 = 0.243$)		
	B	**95% CI**	***p*-Value**	***B***	**95% CI**	***p*-Value**
Participant gender	0.09	−0.27–0.45	0.62	0.18	−0.32–0.67	0.48
Patient gender	−0.24	−0.64–0.17	0.25	−0.19	−2.06–1.68	0.84
Domestic work	0.21	0.07–0.35	0.004	−0.05	−0.28–0.18	0.67
Patient's accuracy	0.36	0.20–0.52	<0.001	0.56	0.30–0.82	<0.001
Domestic work × patient gender				0.42	0.13–0.71	0.005
Accuracy × patient gender				−0.26	−0.60–0.07	0.12
Participant gender × patient gender				−0.30	−1.01–0.41	0.41

Gender of participant and gender of patient are dummy coded with 1 = woman, 0 = man. All other ratings were made on seven-point scales where future work ability was rated between *1 = Very poor* to *7 = Very good*, perceived accuracy of the patient's self-assessed work ability was rated between *1 = Not at all accurate* to *7 = Very accurate*, and perceived amount of domestic work was rated between *1 = Very little* to *7 = Very much*.
Unstandardized regression coefficients, 95% confidence intervals (CI) and *p*-values are presented.

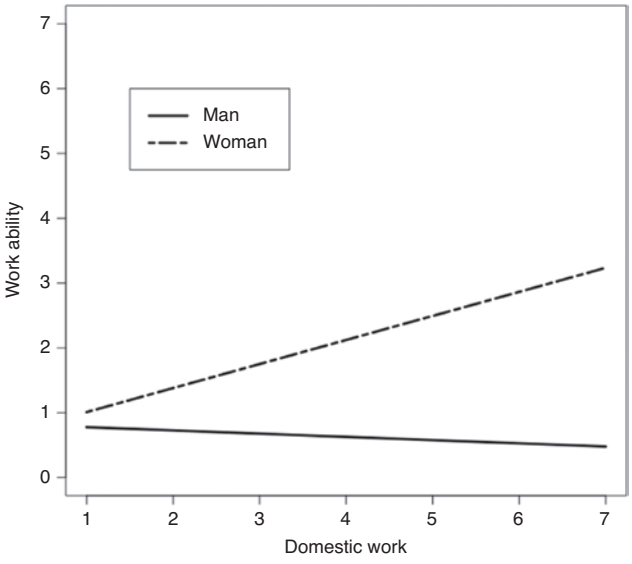

Fig. 1: Influence of perceived level of domestic work on work ability assessments for female and male patients, respectively.
Domestic work (X-axis) and work ability (Y-axis) were assessed on seven-point scales where future work ability was rated between *1 = Very poor* to *7 = Very good*, and perceived amount of domestic work was rated between *1 = Very little* to *7 = Very much*. Estimates are based on the full hierarchical regression model presented above.

decided the level of the patient's self-assessment ability. In the stimuli material, the participants were informed that the patient had a "strong belief in their own work ability". This statement was differently evaluated depending on whether the patient was a man or a woman where male patients were seen as more accurate in assessing their own work ability, than were female patients. This is in line with previous research that shows that men are taken more seriously and that women express that they have to convince their physicians that they are credible [22, 23]. One explanation to why males may be perceived of as better at assessing their own work ability that has been suggested in previous research is that they are perceived as better at describing their symptoms more clearly and demanding [8]. In the present research, there were no differences between the vignettes with respect to how the symptoms were described. Hence, at least in our research we must draw the conclusion that the difference in accuracy perceived by the participants is not due to how the symptoms are described, but rather to stereotypes about women and men. This being said, we did not find support that increased self-assessment accuracy interacted with patient gender in explainihng work ability. Instead, perceptions of the patients' self-assessment accuracy had an impact on work ability assessment for both female and male patients.

Second, women were perceived as doing more domestic work. This resonates with the societal stereotype of women as caregivers and home-makers [9, 14] and it might also reflect real-world segregation between women and men in domestic work. However, the interpretations of these results can be reflected upon. We found that women who were perceived to do more domestic work were considered as having higher work ability. This means that, women who conform to gender stereotypical norms [10] were judged to have higher work ability. A consequence may be that they actually are expected to work more than men, which from a medical perspective instead could result in more stress and ill-health in itself. However, given that women do perform the majority of household chores, it is not

strange that the participants thought that the female patient was also doing more household work than the male. A reasonable conclusion is that anyone who is able to perform domestic work should also be able to perform paid work. Thus, such a conclusion is completely logical. However, the rationale is still affected by gender in the sense that this mechanism only applied to the female patient. That is, only for women is it beneficial to perform domestic work for their perceived work ability. Given that all information about the female and male patient was identical, an unbiased decision process would yield the same results. In relation, recent research show that what the patient perceived that their physician thinks about their work ability is a very strong predictor of their own work ability [29–33] which in turn has been shown to be the most important predictor of actual return to work [24].

4.1 Practical implications and future research

Working with written scenarios of patient descriptions may be a less externally valid method in comparison to having real patients or doing archival studies. However, a written scenario can isolate a gender bias because other factors can be controlled for.

The main outcome variable did not differ between women and men. However, gender stereotypes influenced other judgments that played a role in the decision process. West and Zimmerman [34] have described how gender stereotypes are construed and reinforced in interactions. Thus, if physicians treat female and male patients differently, gender biases may be reinforced in this process. Further, because physicians are authorities in relation to a patient, their attitudes and perceptions of the patient will affect the patients' self-perception and motivation [35]. For instance, pain patients are in a vulnerable position, and individuals in vulnerable situations are more sensitive to gender stereotypes and adapt more readily to them [36]. One way that physicians' stereotypes could affect the patient is through the questions that they ask. Hamberg and Risberg [8] showed that female patients were asked more questions about family life than male patients. Such questions could activate the female stereotype among women and lead women to associate themselves more strongly with domestic roles. Moreover, such an association may also decrease attachment to work life and lower their thresholds for seeking sick leave [37]. Thus, training medical students in equal treatments of women and men in the diagnosing process is of utmost importance.

5 Limitations

Some limitations are worth noting. First, this study utilized a student sample and the results do not necessarily translate into professional physicians. Further, the study was completed in a Swedish context with medical students answering the questionnaire. Thus, it could be questioned whether the results are generalizable to working physicians and other cultures as well. We believe they are, mainly because we found these effects in a country high on gender equality [38] and because younger people have less sexist attitudes than older [39]. Thus, it might be hypothesized that the effects could be larger if sexist attitudes are stronger. It is also important to note that the students enter medical school with existing stereotypes about women and men that could either be augmented or attenuated throughout their training. Importantly, the stereotypes are not likely to be diminished without proper education. This is something that the educational system must take into consideration.

In the present study, we did not define domestic work. It is very plausible that we would have received different results if we had highlighted that domestic work also could include chores that are traditionally coded as masculine (e.g. fixing things around the house [40]). Still, it is important to remember that stereotypes are to some extent averages, but when applied to specific individuals they may lead to biased judgments.

Future research should include other factors that could be important from an equality perspective, such as ethnicity, sexual orientation and age [41]. For example, it could be hypothesized that gender stereotypes are more salient in heterosexual patients and when children live in the household. Besides, different types of diagnoses could be investigated with similar methods. In relation, other socio-demographic characteristics of the patient should be more closely investigated, such as varying education level or income, in the vignette. Thus, it should be tested if gender stereotypes are more or less important in different settings. Although important, such a design would considerably increase the needed sample size.

Stereotypes also operate more strongly under ambiguity and cognitive load [17]. Therefore, it would be interesting to see whether gender differences in assessment increase if physicians are stressed and under pressure.

It is important to note that our independent factors only explained small part of the variation of work ability judgements. The full regression analysis explained about 24% of the variation in work ability judgements. This means that other factors, not included in our model, stood for about 76% of the variation.

6 Conclusions

This study investigated gender stereotypes in relation to patients on sick-leave. It showed that patient gender matters for to-be physicians' estimates of future work ability. Male patients were seen as better at estimating their own work ability, which in turn strongly predicted physicians' estimates. Besides, amount of domestic work predicted work ability among women but not among men.

Authors' statements

Research funding: This work was supported by the AFA Insurance under Grant 140379.

Conflict of interest: The authors declare no conflicts of interest.

Informed consent: All participants gave informed consent before completing the questionnaire.

Ethical approval: The study was reviewed and approved by the Regional Ethics Board at the University of Gothenburg, number 1120-2016.

References

[1] Breivik H, Collett B, Ventafridda V, Cohen R, Gallacher D. Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment. Eur J Pain 2006;10:287–333.

[2] Bergman S, Herrstrom P, Hogstrom K, Petersson IF, Svensson B, Jacobsson LT. Chronic musculoskeletal pain, prevalence rates, and sociodemographic associations in a Swedish population study. J Rheumatol 2001;28:1369–77.

[3] Wijnhoven HA, de Vet HC, Picavet HS. Prevalence of musculoskeletal disorders is systematically higher in women than in men. Clin J Pain 2006;22:717–24.

[4] Stalnacke BM, Haukenes I, Lehti A, Wiklund AF, Wiklund M, Hammarstrom A. Is there a gender bias in recommendations for further rehabilitation in primary care of patients with chronic pain after an interdisciplinary team assessment? J Rehabil Med 2015;47:365–71.

[5] Steensma TD, Kreukels BP, de Vries AL, Cohen-Kettenis PT. Gender identity development in adolescence. Horm Behav 2013;64:288–97.

[6] Emslie C, Hunt K, Macintyre S. Problematizing gender, work and health: the relationship between gender, occupational grade, working conditions and minor morbidity in full-time bank employees. Soc Sci Med 1999;48:33–48.

[7] Courtenay WH. Constructions of masculinity and their influence on men's well-being: a theory of gender and health. Soc Sci Med 2000;50:1385–401.

[8] Hamberg K, Risberg G, Johansson EE, Westman G. Gender bias in physicians' management of neck pain: a study of the answers in a Swedish national examination. J Women's Health Gender Based Med 2002;11:653–66.

[9] Ellemers N. Gender stereotypes. Ann Rev Psychol 2018;69:275–98.

[10] Heilman ME. Description and prescription: how gender stereotypes prevent women's ascent up the organizational ladder. J Social Issues 2001;57:657–74.

[11] Heilman ME, Okimoto TG. Why are women penalized for success at male tasks?: The implied communality deficit. J Appl Psychol 2007;92:81–92.

[12] Heilman ME, Wallen AS. Wimpy and undeserving of respect: penalties for men's gender-inconsistent success. J Exp Soc Psychol 2010;46:664–7.

[13] Ridgeway CL. Framed before we know it: how gender shapes social relations. Gender Soc 2008;23:145–60.

[14] Wood W, Eagly AH. Gender. In: Fiske ST, Gilbert DT, Lindzey G, editors. Handbook of Social Psychology. edn. New York: John Wiley & Sons, Inc., 2010.

[15] Eagly AH, Steffen VJ. Gender stereotypes stem from the distribution of women and men into social roles. J Pers Soc Psychol 1984;46:735–54.

[16] Eagly AH, Wood W. Social role theory. In: van Lange PAM, Kruglanski AW, Higgins ET, editors. Handbook of Theories in Social Psychology. edn. Thousand Oaks, CA: Sage Publications, 2011:458–76.

[17] Fiske S, Lin C, Neuberg SL. The continuum model: ten years later. In: Chaiken S, Tropea P, editors. Dual-process Theories in Social Psychology. edn. New York: Guilford, 1999:231–54.

[18] Stigmar K, Grahn B, Ekdahl C. Work ability – experiences and perceptions among physicians. Disabil Rehabil 2010;32:1780–9.

[19] Bernstein B, Kane R. Physicians' attitudes toward female patients. Med Care 1981;19:600–8.

[20] Norrmén G, Svärdsudd K, Andersson DKG. How primary health care physicians make sick listing decisions: the impact of medical factors and functioning. BMC Fam Pract 2008;9:1–9.

[21] Ahlgren C, Hammarstrom A. Back to work? Gendered experiences of rehabilitation. Scand J Public Health 2000;28:88–94.

[22] Colliver JA, Marcy ML, Travis TA, Robbs RS. The interaction of student gender and standardized-patient gender on a performance-based examination of clinical competence. Acad Med 1991;66(9 Suppl):S31–3.

[23] Werner A, Malterud K. It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. Soc Sci Med 2003;57:1409–19.

[24] Claréus B, Renström EA. Beliefs about physician's opinions regarding return to work volition and ability explain patients' return to work expectancy. J Pain Res (forthcoming) 2018.

[25] Rosenthal R, Jacobson L. Pygmalion in the classroom. Urban Rev 1968;3:16–20.

[26] Wahlin C, Ekberg K, Persson J, Bernfort L, Oberg B. Association between clinical and work-related interventions and return-to-work for patients with musculoskeletal or mental disorders. J Rehabil Med 2012;44:355–62.

[27] Furman G, Colliver JA, Galofre A. Effects of student gender and standardized-patient gender in a single case using a male and a female standardized patient. Acad Med 1993;68:301–3.
[28] Green SB. How many subjects does it take to do a regression analysis. Multivar Behav Res 1991;26:499–510.
[29] Bialosky JE, Bishop MD, Cleland JA. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. Phys Ther 2010;90:1345–55.
[30] Flood AB, Lorence DP, Ding J, McPherson K, Black NA. The role of expectations in patients' reports of post-operative outcomes and improvement following therapy. Med Care 1993;31:1043–56.
[31] Berglind H, Gerner U. Motivation and return to work among the long-term sicklisted: an action theory perspective. Disabil Rehabil 2002;24:719–26.
[32] Franche RL, Krause N. Readiness for return to work following injury or illness: conceptualizing the interpersonal impact of health care, workplace, and insurance factors. J Occup Rehabil 2002;12:233–56.
[33] Darlow B. Beliefs about back pain: the confluence of client, clinician and community. Int J Osteopath Med 2016;20:53–61.
[34] West C, Zimmerman DH. Doing gender. Gender Soc 1987;1: 125–51.
[35] Mitchell TR, Daniels D. Motivation. In: Borman WC, Igen RJ, Kimoski RJ, editors. Handbook of Psychology. edn. Hoboken, NJ: John Wiley & Sons, Inc., 2003.
[36] Schimel J, Simon L, Greenberg J, Pyszczynski T, Solomon S, Waxmonsky J, Arndt J. Stereotypes and terror management: evidence that mortality salience enhances stereotypic thinking and preferences. J Pers Soc Psychol 1999;77:905–26.
[37] Angelov N, Johansson P, Lindahl E. Kvinnors större föräldraansvar och högre sjukfrånvaro (Women's higher responsibility as parents and women's higher sick leave). In: Edited by (IFAU) Ifa-ouu, vol. 2013:7. Uppsala; 2013.
[38] Global Gender Gap Report. Retrieved from: http://reports.weforum.org/global-gender-gap-report-2017/results-and-analysis/). [Retrieved from: http://reports.weforum.org/global-gender-gap-report-2016/top-ten/].
[39] Akrami N, Ekehammar B, Yang-Wallentin F. Personality and social psychology factors explaining sexism. J Individ Differ 2011;32:153–60.
[40] Diekman AB, Eagly AH. Stereotypes as dynamic constructs: women and men of the past, present, and future. Pers Soc Psychol B 2000;26:1171–88.
[41] Eckstrand KL, Eliason J, St Cloud T, Potter J. The priority of intersectionality in academic medicine. Acad Med 2016;91:904–7.