(Self-)Evaluation of computer competence: How gender matters

Monika Sieverding *, Sabine C. Koch

University of Heidelberg, Department of Psychology, Hauptstr. 47-51, 69117 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 9 August 2008
Received in revised form 6 November 2008
Accepted 17 November 2008
Available online xxxx

Keywords:
Gender studies
Media in education
Computer competence
Attribution of success
Self-evaluation

ABSTRACT

Is the negative stereotype of women with regard to computer competence still exerting power in our society? In this study, 206 participants observed a target person (either a woman or a man) on a video who was about to solve a complex computer task. Participants had to estimate whether the target person was successful on this task in a limited amount of time. After they had received the information that the target person had solved the task successfully in the required time period, and that the person’s performance was above average, they were asked to provide a reason for the success (luck vs. skill attribution) and to evaluate the general computer competence of the target. Then, participants had to evaluate their own (hypothetical) computer competence in comparison to the target. Results suggest that for the direct evaluation of the target persons and for the causal attribution of success, no systematic gender-related biases occurred. In the self-ratings of participants, however, findings showed that (a) women judged their computer competence to be lower than did men, and (b) both women and men judged their own hypothetical performance in the computer-related task to be relatively higher when comparing it to the identically scripted performance of a woman vs. a man.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Women, computer competence, and attribution

In a society that is increasingly dependent on information technology, the effective use of computers has become a key qualification for professional success and advancement (Cooper, 2006; Cooper & Weaver, 2003; Levine & Donitsa-Schmidt, 1997; Wasserman & Richmond-Abbott, 2005). A number of studies have investigated the role of gender in computer-related attitudes and in different aspects of computer use, and this research has demonstrated that a gender gap in computer use still exists across all age groups and cultures (Colley & Comber, 2003; Imhof, Vollmeyer, & Beierlein, 2007; Shashaani, 1997). A meta-analysis by Whiteley (1997) conveyed that men and boys exhibited higher self-efficacy – a measure of task-related self-confidence – with respect to computer use than women and girls. Several reasons for the decreased computer self-efficacy beliefs of women have been discussed; for example, the lack of female role models who use computers (Mark & Roman, 2002), gender-specificity of the vast majority of computer software (Cooper, 2006), or gender-specific differences in support from parents to teachers (Busch, 1996). Moreover, unfavourable computer attitudes (Anderson, Lankshear, Timms, & Courtney, 2008; Durndell, Haag, & Laithwaite, 2000; Shashaani, 1997) and computer anxiety (Chua, Chen, & Wong, 1999) may play important roles in contributing to the fact that women do not live up to their potential in the computer domain. Other possible causes could be gender stereotypes and prejudice against girls and women with regard to their computer skills (Cooper, 2006; Cooper & Weaver, 2003; Smith, Morgan, & White, 2005), and/or unfavourable attribution patterns (Nelson & Cooper, 1997; Rozell & Gardner, 1999).

In a classical study Deaux and Emswiller (1974), showed that the successful performance of women on a masculine task was evaluated less favourably than the comparable performance of men on the same task. They further found that the successful performance of a woman was perceived to be more the result of luck (external attribution), whereas the equivalent performance of a man was attributed more to skill (internal attribution). Studies that tried to replicate these findings have yielded heterogeneous results (Swim & Sanna, 1996). With regard to external evaluations and attributions of performance, although a gender bias could not consistently be found, some authors found gender differences in self-attributions. For example Dickhäuser and Stiensmeyer-Pelster (2002), found more unfavourable computer-related attributions in explaining one’s own failure at the computer for women as compared to men. Koch, Müller, and Sieverding (2008) experimentally induced a computer-related failure: Due to a faulty USB-memory stick, task completion was not possible. They found that...
women who were initially confronted with a negative stereotype about women and computers attributed the failure more internally (to their own inability) when compared to women from a control group to men, both of whom attributed the failure more externally (to the faulty technical equipment).

The present study aimed to investigate the following questions: Is the performance of women on computers evaluated less favourably than the performance of men? More specifically, is there still an attributional gender bias as observed by Deaux and Emswiller (1974)? Do women underestimate their own performances? And what happens if the standard of comparison for the self-evaluation is the performance of a male vs. a female target?

Prior research showed that women frequently underestimate their achievements, especially in masculine-stereotyped tasks or domains. In a study by Ehrlinger and Dunning (2003), female college students performed equivalently to male students on a science test, yet they underestimated their performances because they thought less of their general scientific reasoning abilities. They subsequently were more likely to turn down the opportunity to participate in a science competition with attractive prizes. In a recent study by Bosak and Sczesny (2008), female managerial students judged themselves as less suitable for a leadership position than their male counterparts. These findings indicate a pervasive gender bias in self-concepts related to performance in masculine-stereotyped tasks and domains. Does this bias extend to the domain of gender and computers – a domain increasingly relevant in professional as well as in private contexts?

1.2. Research questions and hypotheses

The aim of this study was to generally investigate whether the computer competence of women would be evaluated less favourably than the computer competence of men with regard to external assessments as well as for self-evaluations. If participants had to predict whether a person is able to solve a complex computer task, would the sex of the stimulus person influence this prediction? After successful performance was indicated to participants, would the success of a woman be evaluated less favourably when compared to the success of a man? We were further interested in the question of whether attributions would reveal a gender bias as was found by Deaux and Emswiller (1974) for judgments of the identically successful performances of a man and a woman (“what is skill for the male is luck for the female”). A third goal of our study was to assess gender differences in self-evaluations of computer competence as a function of the sex of the comparison object.

We expected that in scripted scenes shown from a videotape there would be a gender difference in external evaluations. Specifically, we expected that (a) participants would predict the computer competence of a female target person to be lower than the computer competence of a male target person; and (b) the successful performance of the female target person would be more attributed to luck, whereas the successful performance of the male target person would be more attributed to ability. Regarding the self-evaluations of computer competence, we expected that in comparison to the target persons, (a) women would rate their computer skills lower than men under identical circumstances; and that (b) women would rate their competences lower when compared to a male target than when compared to a female target. We tested these hypotheses with a quasi-experimental 2 × 2 design with sex of participant and sex of target as independent variables. Prognosis of the target person’s expected performance at baseline, attribution of the target person’s successful computer-related performance (skill vs. luck), and self-evaluation of competence in comparison to the target person were the dependent variables.

2. Method

2.1. Sample

The participants were 206 students at the Free University of Berlin (99 men and 107 women; 113 psychology students and 93 students from other disciplines); the mean age was 23.7 years (SD = 4.2). Each participant was randomly assigned to one of the two experimental conditions: The first condition required the evaluation of a male target, and the second condition required the evaluation of a female target. Approximately the same number of men and women participated in each of the conditions: 49 men and 53 women evaluated the male target, whilst 50 men and 54 women evaluated the female target. Participants were offered a cup of coffee or tea and a piece of cake as a small reward for participation in the study.

2.2. Procedure

Participants took part in small groups of 4–8 persons. Two female experimenters conducted the study. First, the following instructions were read (in German): This study is investigating the ability to predict and to evaluate the computer performance of other people. You will see a video recording of a student who had to solve a task at a personal computer. The task consisted of (a) the correction of a faulty text, (b) the search for and installation of the correct printer–driver, and (c) the printing of the corrected text without mistakes. The students who participated in this study had exactly 30 min to solve the task. If they did not successfully solve the task within this time, the trial was interrupted. Sixty percent of the students solved the task. The main difficulty was finding and installing the correct printer driver. You will now be asked to predict the performance of the student on the video recording and to evaluate his or her performance later.

The participants then watched the video recording (2 min). They saw a target person – of about their age – entering a room and receiving instructions from a female experimenter. In the video, the target person was told by a female experimenter that the task consisted of (a) the correction of a faulty text, (b) the search for and installation of the correct printer–driver, and (c) the printing of the corrected text without mistakes. The target person was informed of the 30 min time limit to solve the task and was told that the trial would be interrupted if they had not solved the task by then. Next, participants saw the experimenter ask the target person whether they had any questions (which was denied by the target persons), and then ask them to sit down at a table with a personal computer and a printer. The target person sat down and began to work at the computer. At this point – after 2 min – the video recording ended.

Following this first sequence, participants were asked to make a guess about whether the performance of the target was successful or not. If they predicted a successful performance, they were asked to indicate the estimated time the target person needed to solve the problem. After the participants had provided their estimates for the expected performance of the target person, they received the feedback...
about the alleged performance of the target person on the monitor. A slide was shown indicating to the participant that the target person had successfully solved the task and that their performance was above average, specifically, that they had solved the task in 15.04 min, whereas the average task solution took 24.02 min. Then, they were asked to evaluate this performance on two scales. Finally, they judged their own hypothetical performance in comparison to the target person, in the case that they would have to perform the same task. At the end of the experiment, participants were carefully debriefed and asked not to talk to others about the experiment.

2.3. Training of targets

Four potential targets (2 male and 2 female students close to the participants’ mean age) were trained to play their roles with verbatim scripts in a single session of 60 min, and rehearsed the scene several times before they were actually recorded. A female student and a male student who were evaluated as equally self-confident and competent in their videotaped behaviour (by a group of students and the two authors of the study) were chosen as target persons.

2.4. Instruments and scales

2.4.1. Predicted performance and time estimation

Participants were asked to guess whether the target person would successfully solve the task in the given time (prediction of performance at baseline). The three response alternatives were: probably yes, probably no, and I do not know. Participants were also asked to estimate the time (in minutes) that the target would need to solve the task (open question with no numbers given as prompts).

2.4.2. Evaluation of the target’s performance

Participants were asked to evaluate the (alleged) above average successful performance of the target person on two 13-point scales that were modified from the study by Deaux and Emswiller (1974). The first scale measured the attribution of the performance on a luck-skill dimension. This bipolar scale was labelled “pure luck” at one end and “pure ability” at the other end with a midpoint label of “both luck and ability in equal parts.” The second scale assessed the participants’ evaluations of the general computer competence of the target person, from “very low,” through “average,” to “very high.”

2.4.3. Self-evaluation of computer competence

Participants were asked to estimate their own hypothetical performance in comparison to the targets’ performance, in the case that they would have performed the same task, on a 13-point-scale from “very much worse,” through “comparable,” to “very much better.”

2.5. Statistical analyses

2.5.1. Performance prediction and time estimation

A Chi-square analysis was conducted on the performance predictions. The estimated solution time was analysed with a $2 \times 2$ ANOVA with sex of target and sex of participant as independent factors.

2.5.2. Performance evaluation of the target

We computed a $2 \times 2$ MANOVA with the independent between-group variables, sex of participant and sex of target, and the dependent variables, attribution of success and evaluation of computer competence.

2.5.3. Self-evaluation of computer competence

Self-evaluation of participants in comparison to the target was analysed with a $2 \times 2$ ANOVA with sex of target and sex of participant as independent factors.

The alpha-level was 0.05. As variables of potential relevance with respect to computer skill level, we controlled age and major, which did not exert an influence.

3. Results

3.1. Predicted performance and time estimation

A comparison of participants’ predictions regarding the performance of targets did not show a significant effect of sex of target, $\chi^2(2, N = 206) = 0.18, p = 0.91$. Regardless of sex of target, the majority of participants expected the target to be successful in solving the task (83% when the target was a woman, and 84% when the target was a man). There further was no significant effect of sex of participant. Female and male participants did not differ significantly in their predicted performance, $\chi^2(2, N = 206) = 1.4, p = 0.51$. Moreover, there was no gender effect for the predicted solution time, neither for sex of target, $F(1, 202) = 0.01, p = 0.98, \eta^2 = 0.00$, nor for sex of participant, $F(1, 202) = 1.85, p = 0.17, \eta^2 = 0.01$. The predicted solution time for the female target was $M = 18.6$ min ($SD = 8.7$), and for the male target, $M = 18.7$ min ($SD = 7.6$). Our first hypothesis, which postulated that the performance predictions would be lower when the target person was female, was therefore, not supported by the data.

3.2. Performance evaluation of the target

The $2 \times 2$ MANOVA revealed no significant effect of sex of participant or sex of target for either evaluation of computer competence of target or attribution of success. The successful performance of the female target was not attributed more to luck than to ability when compared to the male target, $F(2, 202) = 1.24, p = 0.27, \eta^2 = 0.01$ for sex of target, and $F(2, 202) = 0.24, p = 0.62, \eta^2 = 0.00$ for sex of participant.
In this study, we investigated gender influences on performance expectations, attribution of success, perceived computer competence, and self-evaluations of computer competence as compared to a successful target. Results suggest no gender effects on performance expectations, attributions, and perceived computer competence of male and female targets. In these aspects, the study did not confirm the hypotheses that women’s abilities would be underestimated with regard to computer tasks and that the success of women would be attributed to luck rather than skill.

Yet, self-evaluations of computer competence depended on the sex of the participant and on the sex of the target. The latter finding speaks to the shifting standard effect (Biernat & Kobrynowicz, 1997) and the double standards effect in general (Carli & Eagly, 1999; Foschi, 2000). When I see my own computer competence in a better light when compared to a woman than when compared to a man, I am implicitly using different standards of comparison for the evaluation of male and female performance. This finding further indicates that the stereotype still exists, but only emerges in the data if it is made self-relevant. An interesting result of this study is that computer self-efficacy is influent by the social context within which individuals evaluate themselves (Cooper, 2006; Inzlicht & Ben-Zeev, 2000). The fact that women showed lower computer self-efficacy than men is in line with former research (Durndell et al., 2000; Koch et al., 2008; Whitley, 1997).

3.3. Self-evaluation of computer competence

The analysis of the self-evaluations of computer competence of the participants in comparison to the target resulted in a significant main effect of sex of participant, $F(2, 202) = 21.24$, $p < 0.001$, $\eta^2 = 0.10$, and sex of target, $F(1, 202) = 4.88$, $p < 0.05$, $\eta^2 = 0.02$ (see Fig. 1). Self-evaluations by male participants were higher than those by female participants, and self-evaluations were higher when the target was a woman (see Table 1). There was no significant sex of participant $\times$ sex of target interactions. Only 7% of the female participants rated their own (hypothetical) performance as comparable to that of the male target person, and only one woman rated her performance as better. The great majority thought that their own performance would be worse (scores of 6 or less), and 25% of the female participants chose the lowest score of 1, which means they rated their own hypothetical performance “very much worse.” The majority of the male participants also rated their hypothetical performance as less successful in comparison to the male target, but 22% thought they would show a comparable performance, and 8% rated their own hypothetical performance as better.

4. Discussion

4.1. Gender effects on self-evaluation

In this study, we investigated gender influences on performance expectations, attribution of success, perceived computer competence, and self-evaluations of computer competence as compared to a successful target. Results suggest no gender effects on performance expectations, attributions, and perceived computer competence of male and female targets. In these aspects, the study did not confirm the hypotheses that women’s abilities would be underestimated with regard to computer tasks and that the success of women would be attributed to luck rather than skill.

Yet, self-evaluations of computer competence depended on the sex of the participant and on the sex of the target. The latter finding speaks to the shifting standard effect (Biernat & Kobrynowicz, 1997) and the double standards effect in general (Carli & Eagly, 1999; Foschi, 2000). When I see my own computer competence in a better light when compared to a woman than when compared to a man, I am implicitly using different standards of comparison for the evaluation of male and female performance. This finding further indicates that the stereotype still exists, but only emerges in the data if it is made self-relevant. An interesting result of this study is that computer self-efficacy is influenced by the social context within which individuals evaluate themselves (Cooper, 2006; Inzlicht & Ben-Zeev, 2000). The fact that women showed lower computer self-efficacy than men is in line with former research (Durndell et al., 2000; Koch et al., 2008; Whitley, 1997).
Research findings on the role of self-efficacy beliefs indicate that these are an important aspect of successful behaviour such as a successful vocational career (Bandura, 1995). The lower self-efficacy beliefs of women regarding their competence in information technologies, and particularly computers, may have negative consequences for their vocational careers and for society in general. A woman who thinks about entering a vocational field where computer competence is necessary will usually be confronted with the fact that a majority of highly competent and successful men are working in this field. According to the findings from this study, negative consequences for women’s computer self-efficacy are to be expected: When women compare their own performance to that of the men in their professional contexts, a resulting high computer self-efficacy in women will be the exception rather than the rule. A practical result of the findings could be the training of computer competence (and other stereotyped competences) in gender homogeneous groups in diverse stages of the educational and training sector. This way, girls and women could acquire a higher baseline computer self-efficacy and would not as readily suffer from the effects of negative comparison to stereotype threat (see for example Crombie & Armstrong (1999)).

4.2. Limitations of the study

Limitations of the study include the stimulus sampling problem (Wells & Windschitl, 1999). In this study, we only had one male and one female stimulus. Since it is absolutely impossible to show the exact same non- and paraverbal behaviours (such as head tilts, pauses, intonation, etc.) in a scripted performance, the internal validity of such experiments is limited. Further, actual gender differences in computer competence of participants were not tested in this study therefore, we cannot make claims about the relation of the reported subjective differences and any objective differences in computer competence in this sample. Future studies on self-evaluations of computer competence should assess computer skills with objective measures and compare self-evaluations with actual skills. Regarding the measures used, self-evaluation was provided on a single scale, thereby compromising reliability. Future studies should include more differentiated measures of self-evaluation. The effect size for the influence of the target as a standard of comparison was $\eta^2 = 0.02$; this is a very small effect and should therefore, not be over-interpreted. Moreover, the sample was limited to university students and in order to generalise to the broader population, a more heterogeneous sample is needed.

4.4. Conclusions and outlook

In sum, the results of Deaux and Emswiller (1974) that performance expectations for women in a masculine task (at the computer) are lower and that attribution patterns are negatively biased were not confirmed. It seems that amongst German students nowadays, negative stereotypes regarding the performance of women at the computer are no longer relevant, if measured directly and if the target persons appear similarly competent. This does not exclude that there may still be such stereotypes in older and more traditional parts of society. Gender-biased competence stereotypes might also be more influential for female targets who conform to traditional gender stereotypes (for example, in their verbal or non-verbal behaviour or in their outfit). In our study, we used target persons who were very similar in their behaviour and in their outer appearance, which actually may have caused lower salience of sex of target and thus, limited causal attributions to sex. Future studies should analyse the evaluation of male and female targets who match traditional gender stereotypes in their outer appearance or in their behaviour. Sczesny, Spreemann, and Stahlberg (2006) have shown in experimental research that the attribution of leadership competence to women and men was influenced by their physical appearance. Persons with a typical feminine appearance were evaluated as less competent than persons with a typical masculine appearance (Sczesny et al., 2006). We assume that similar effects might be found in the evaluation of the computer competence of women (and men).

Generally, however, self-evaluations of computer competence did depend on the standard of comparison, that is, the sex of the reference person (main effect of sex of target) independent of participant’s sex. Moreover, women rated their computer competence as worse than men (main effect of sex of target) independent of sex of the reference person.

The small effect for sex of target as a standard of comparison and the null findings of the first hypothesis are generally indicative of a positive development. Much has changed for the better (i.e., more gender-equal) in the gender-related perception of self and others in the computer domain. However, there is still some work left to do as indicated by the more indirect measure of self-evaluation: Men estimated their computer competence more highly than women, and both men and women estimated their own competences more highly in comparison with a female target than with a male target. Throughout the school experience, since there are usually more men than women in computer classes, the mechanisms found in this study can cause unwanted social comparison effects and a threatening environment for girls and women. Thus, there is still some way to go before men and women will be viewed and will view themselves as equally competent in the computer domain.

Acknowledgements

Monika Sieverding and Sabine C. Koch, Department of Psychology, University of Heidelberg, Germany. We wish to thank Eva Horvath, Tanja Nord, Maja von Strempel, Daniela Vogt and Christine Lauterbach for their support in preparing and conducting the experiment.

References


